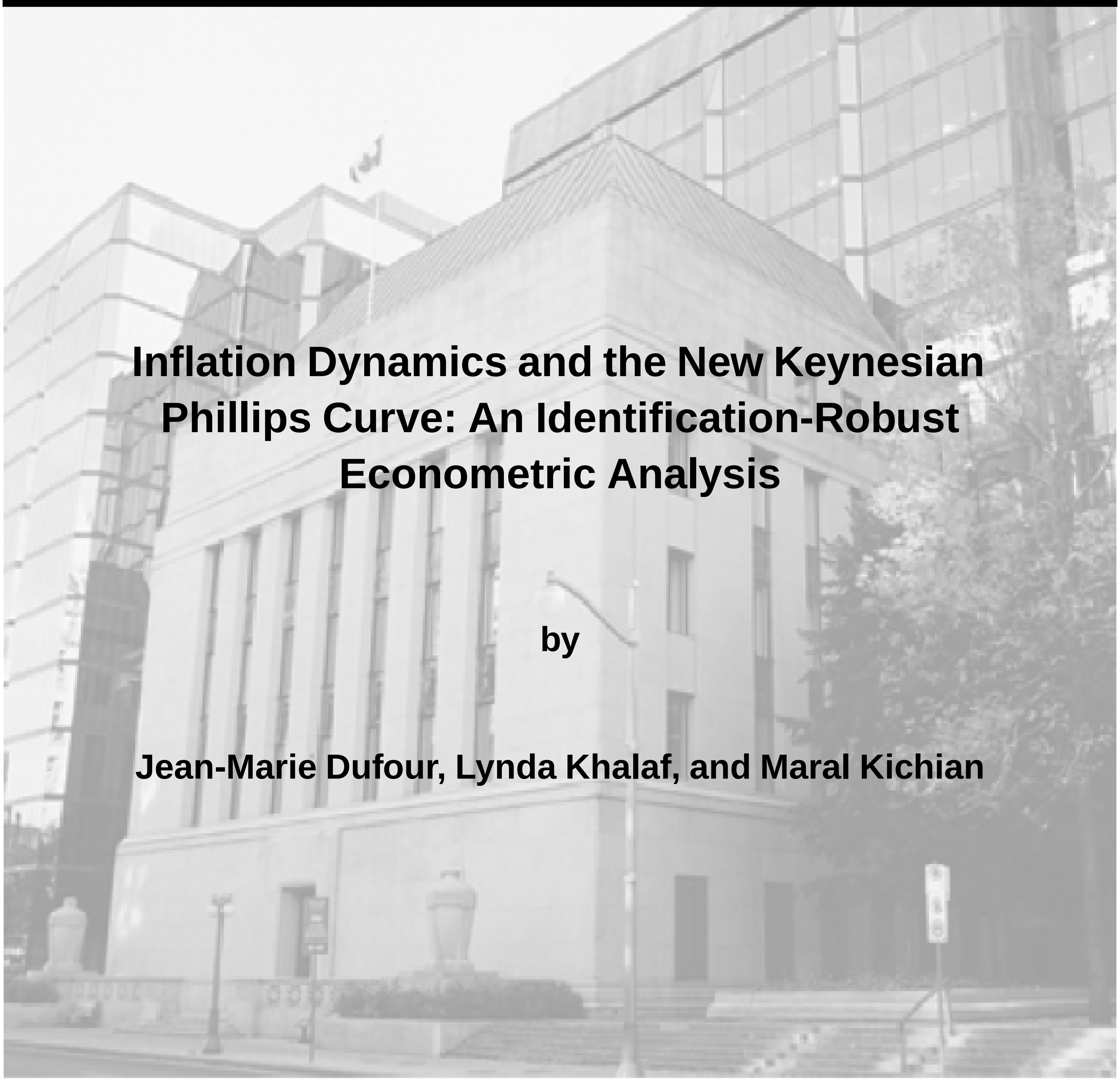
Bank of Canada Banque du Canada

Working Paper 2005-27 / Document de travail 2005-27



Inflation Dynamics and the New Keynesian Phillips Curve: An Identification-Robust Econometric Analysis

by

Jean-Marie Dufour, Lynda Khalaf, and Maral Kichian

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2005-27

October 2005

Inflation Dynamics and the New Keynesian Phillips Curve: An Identification-Robust Econometric Analysis

by

Jean-Marie Dufour,[1] Lynda Khalaf,[2] and Maral Kichian[3]

[1]Canada Research Chair Holder (Econometrics)
CIRANO, CIREQ
Département de sciences économiques
Université de Montréal, Quebec, Canada H3C 3J7
jean.marie.dufour@umontreal.ca

[2]Canada Research Chair Holder (Environment)
Département d'économique and GREEN
Université Laval, Quebec, Canada G1K 7P4
and
CIREQ
Université de Montréal, Quebec, Canada H3C 3J7
lynda.khalaf@ecn.ulaval.ca

[3]Research Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
mkichian@bankofcanada.ca

Contents

Acknowledgements

We would like to thank Florian Pelgrin and two anonymous referees for useful comments, and Wendy Chan and Amanda Armstrong for research assistance. This work was supported by the Canada Research Chair Program (Econometrics, Université de Montréal, and Environment, Université Laval), the Institut de Finance mathématique de Montréal (IFM2), the Alexander-von-Humboldt Foundation (Germany), the Canadian Network of Centres of Excellence (program on Mathematics of Information Technology and Complex Systems [MITACS]), the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Quebec), the Fonds de recherche sur la nature et les technologies (Quebec), and the Chair on the Economics of Electric Energy (Université Laval).

Abstract

The authors use identification-robust methods to assess the empirical adequacy of a New Keynesian Phillips curve (NKPC) equation. They focus on Galí and Gertler's (1999) specification, for both U.S. and Canadian data. Two variants of the model are studied: one based on a rational-expectations assumption, and a modification to the latter that uses survey data on inflation expectations. The results based on these two specifications exhibit sharp differences concerning: (i) identification difficulties, (ii) backward-looking behaviour, and (iii) the frequency of price adjustment. Overall, the authors find that there is some support for the hybrid NKPC for the United States, whereas the model is not suited to Canada. Their findings underscore the need for employing identification-robust inference methods in the estimation of expectations-based dynamic macroeconomic relations.

JEL classification: C13, C52, E31
Bank classification: Econometric and statistical methods; Inflation and prices

Résumé

Pour éprouver la validité empirique d'une équation fondée sur la nouvelle courbe de Phillips keynésienne, les auteurs adoptent des méthodes d'inférence qui permettent de surmonter le problème de l'identification. Ils appliquent la spécification proposée par Galí et Gertler (1999) à l'étude des données américaines et canadiennes. Deux variantes du modèle retiennent leur attention. La première repose sur l'hypothèse des anticipations rationnelles, et la deuxième est basée sur une version modifiée de celle-ci qui met à profit des mesures de l'inflation attendue tirées d'enquêtes. Les résultats obtenus à l'aide de ces deux variantes sont nettement différents en ce qui concerne : i) le problème de l'identification; ii) le comportement adaptatif; et iii) la fréquence de rajustement des prix. Les auteurs concluent que la nouvelle courbe de Phillips keynésienne hybride décrit assez bien la dynamique de l'inflation dans le cas des États-Unis mais pas dans celui du Canada. Leurs observations confirment l'utilité de méthodes d'inférence robustes sur le plan de l'identification pour l'estimation de relations macroéconomiques dynamiques faisant intervenir des anticipations.

Classification JEL : C13, C52, E31
Classification de la Banque : Méthodes économétriques et statistiques; Inflation et prix

1. Introduction

A standard feature of macroeconomic policy models is an equation that describes the evolution of inflation. Nowadays, this process is typically modelled as a hybrid New Keynesian Phillips curve (NKPC). This specification results from the efforts of recent years to model the short-run dynamics of inflation starting from optimization principles; see, for example, Woodford (2003) and the references cited therein. In its basic form, the NKPC stipulates that inflation at time t is a function of expected future inflation and the current output gap. With its clearly elucidated theoretical foundations, the NKPC possesses a straightforward structural interpretation and therefore presents, in principle, a strong theoretical advantage to the traditional reduced-form Phillips curve (which is justified only statistically).

Given the statistical failure of the basic NKPC formulation when confronted with the data, however, the curve has evolved into its more empirically viable hybrid form. In particular, it has been noted that: (i) adding a lagged inflation term to the above model (referred to as a *hybrid NKPC*) corrects the signs of estimated coefficients (Fuhrer and Moore 1995; Roberts 1997; Fuhrer 1997), and (ii) using a measure of real marginal cost derived from a given production function instead of the output gap provides a better statistical fit according to generalized method of moments (GMM)-based estimates and tests (Galí and Gertler 1999; Galí, Gertler, and Lopez-Salido 2001). Yet the question of which production function (i.e., which marginal cost measure) is empirically preferable is not yet resolved, because the choice for the marginal cost proxy seems to affect evidence on the weight of the backward-looking term (Gagnon and Khan 2005). In addition, there are different theoretical ways of incorporating backward-looking behaviour in the curve, and which yield differing outcomes (Fuhrer and Moore 1995; Galí and Gertler 1999; Eichenbaum and Fisher 2004).[1]

Discriminating between competing alternatives calls for reliable econometric methods. Full-information models are typically non-linear and heavily parameterized.[2] Therefore, in practice, these models are often estimated by applying standard limited-information (LI) instrumental variable (IV) methods to first-order conditions of interest. Indeed, the popularity of NKPC models stems in large part from studies such as Galí and Gertler (1999) and Galí, Gertler, and Lopez-Salido (2001), who find empirical support for their version of the curve using GMM, and the fact that the model is not rejected according to Hansen's J-test.

But even as the popularity and usage of the NKPC has grown, criticisms have been raised with respect to its empirical identifiability with IV-based methods. The main issue is that IV methods such as GMM are not immune to the presence of weak instruments (Dufour 1997, 2003; Staiger and Stock 1997; Wang and Zivot 1998; Zivot, Startz, and Nelson 1998; Stock and Wright 2000; Dufour and Jasiak 2001; Stock, Wright, and Yogo 2002; Kleibergen 2002; Khalaf and Kichian 2004, 2005; Dufour and Khalaf 2003a,b; Dufour and Taamouti 2003a,b, 2004, 2005). These studies demonstrate that standard asymptotic procedures (which *impose identification away* without correcting for local-almost-identification) are fundamentally flawed and lead to spurious overrejections, even with fairly large sample sizes. In particular, the following fundamental problems occur: in models that may not be

[1] For example, Galí and Gertler (1999) appeal to the assumption that a proportion of firms never reoptimize, but set their prices using a rule-of-thumb method; Eichenbaum and Fisher (2004) use dynamic indexing instead.

[2] In this literature, some of the parameters are typically calibrated, while others are estimated.

identified over all the parameter space, (i) standard t-type tests have significance levels that may deviate arbitrarily from their nominal levels, since it is not possible to bound their null distributions, and (ii) Wald-type confidence intervals—of the form: estimate $\pm$ (asymptotic standard error) $\times$ (asymptotic critical point)—have dramatically poor coverage irrespective of their nominal level, because they are bounded by construction (Dufour 1997).[3]

To circumvent difficulties associated with weak instruments, the above-cited recent work on IV-based inference has focused on two main directions (see the surveys of Dufour 2003 and Stock, Wright, and Yogo 2002): (i) refinements in asymptotic analysis that hold whether instruments are weak or not (e.g., Staiger and Stock 1997; Wang and Zivot 1998; Stock and Wright 2000; Kleibergen 2002; Moreira 2003b), and (ii) finite-sample-justified procedures based on proper pivots; i.e., statistics whose null distributions do not depend on nuisance parameters or can be bounded by nuisance-parameter-free distributions (*boundedly pivotal functions*) (Dufour 1997; Dufour and Jasiak 2001; Dufour and Khalaf 2002; and Dufour and Taamouti 2003a,b, 2004, 2005). The latter include methods based on Anderson and Rubin's (1949, AR) *pivotal* F-statistic that allow *unbounded* confidence sets.

Identification difficulties have led several authors to re-examine the NKPC models, particularly the Galí and Gertler NKPC specification. Linde (2001) performs a small-scale simulation study based on a Galí-Gertler-type model and documents the superiority of full-information maximum likelihood (FIML) relative to GMM. In particular, GMM estimates appear sensitive to parameter calibrations. Ma (2002) applies the asymptotic methods proposed in Stock and Wright (2000) to Galí and Gertler's NKPC in view of getting confidence sets that account for the presence of weak instruments. These sets prove to be much too large to be informative, which suggests that the parameters of the curve are indeed not well identified. Nason and Smith (2003) study the identification issue of the NKPC in limited-information contexts analytically, solving the Phillips curve difference equation. They show that typical GMM estimations of such curves have parameters that are not identifiable (or nearly so), and that FIML can make identification easier. Applications to U.S. data yield GMM estimates that are comparable to the values obtained by Galí and Gertler (1999). In contrast, Nason and Smith's FIML estimates (which they feel are more reliable) point to a greater role for backward-looking behaviour. For Canada, Nason and Smith report that the NKPC is poorly identified, whether GMM or FIML estimation is used. Finally, Fuhrer and Olivei (2004) consider improved GMM estimation, where the instrumentation stage takes the constraints implied by the structure formally into consideration. They demonstrate (via a Monte Carlo simulation) the superiority of their approach. In addition, Fuhrer and Olivei estimate an inflation equation using U.S. data, and obtain a large forward-looking component with conventional GMM, but a much lower value for this parameter with "optimal" GMM and maximum likelihood.

In this paper, we reconsider the problem of estimating inflation dynamics, in view of these

[3]Poor coverage (which implies that the data are uninformative about the parameter in question) is not really due to dramatically large estimated standard errors, or even to poorly approximated cut-off points. The problems stem from the method of constructing the confidence set as an interval that is by definition "bounded." Any valid method for the construction of confidence sets should allow for possibly unbounded outcomes, when the admissible set of parameter values is unbounded (as when parameters are not identifiable on a subset of the parameter space). In this case, a bounded confidence set would inevitably "rule out" plausible parameter sets, with obvious implications for coverage.

recent econometric findings. Our aim is to derive identification-robust estimators and confidence sets. Identification-robust procedures should lead to uninformative (e.g., unbounded) confidence sets when the parameters considered are not identified (Dufour 1997). We focus on two types of procedures: the AR procedure and a method proposed by Kleibergen 2002. The AR procedure is particularly appropriate from the viewpoint of validating a structural model, because it is robust not only to weak instruments, but also to missing instruments and more generally to the formulation of a model for endogenous explanatory variables (Dufour 2003; Dufour and Taamouti 2003 a,b). A drawback of the AR procedure is that it leads to the inclusion of a potentially large number of additional regressors (identifying instruments), and hence to a reduction in degrees of freedom, which can affect the test power in finite samples. To assess sensitivity to this type of effect, we apply a method proposed by Kleibergen (2002), which may yield power gains by reducing the number of "effective" regressors (although at the expense of some robustness).[4]

Our applications study U.S. and Canadian data using: (i) the benchmark hybrid NKPC of Galí and Gertler, which uses a rational-expectations assumption, and (ii) a modification to the latter that uses survey-based measures of expected inflation. Our analysis allows us to compare and contrast both variants of the model; this is relevant because available studies imply that the specification of the expectations variable matters empirically. For instance, Galí and Gertler 1999 suggest that, when the model is conditional on labour costs, under rational expectations, additional lags of inflation are no longer needed. In contrast, Roberts 2001 argues that those results are likely sensitive to the particular specification of labour costs, and that the need to include additional lags could reflect the fact that expectations are not rational; see also Roberts 1997, 1998. Our results reveal very sharp differences with both specifications for the United States and Canada.

In section 2, we reproduce Galí and Gertler's (1999) NKPC hybrid specification. In section 3, we describe the specific models and the methodology used in this paper. Section 4 discusses our empirical results, and section 5 concludes. Appendixes A and B provide details and a formal treatment of the statistical procedures that we apply.

2. Galí and Gertler's Hybrid NKPC Model

In Galí and Gertler's hybrid specification, firms evolve in a monopolistically competitive environment and cannot adjust their prices at all times. A Calvo-type assumption is used to represent the fact that a proportion of firms, θ, do not adjust their prices in period t. In addition, it is assumed that some of the firms do not optimize but use a rule of thumb when setting their prices. The proportion of such firms (referred to as backward-looking price-setters) is given by ω. In such an environment, profit maximization and rational expectations lead to the following hybrid NKPC equation for inflation (π_t):

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

with

$$\pi_{t+1} = E_t \pi_{t+1} + \upsilon_{t+1}, \tag{2}$$

[4] For further discussion of this issue, see Dufour and Taamouti (2003a,b).

and for

$$\lambda = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\theta + \omega - \omega\theta + \omega\beta\theta} \tag{3}$$

$$\gamma_f = \frac{\beta\theta}{\theta + \omega - \omega\theta + \omega\beta\theta}, \gamma_b = \frac{\omega}{\theta + \omega - \omega\theta + \omega\beta\theta}. \tag{4}$$

$E_t\pi_{t+1}$ is expected inflation at time t, s_t represents real marginal costs (expressed as percentage deviation with respect to its steady-state value), and υ_t is unexpected inflation. The parameter γ_f determines the forward-looking component of inflation, and γ_b its backward-looking part; β is the subjective discount rate.

Galí and Gertler rewrite the above NKPC model in terms of orthogonality conditions. Two different normalizations are used for this purpose.[5] The first, orthogonality specification (1), is given by:

$$E_t\left\{(\pi_t - \lambda s_t - \gamma_f \pi_{t+1} - \gamma_b \pi_{t-1})\, z_t\right\} = 0, \tag{5}$$

and the other, orthogonality specification (2), is written as:

$$E_t\left\{(\phi\pi_t - \phi\lambda s_t - \phi\gamma_f \pi_{t+1} - \phi\gamma_b \pi_{t-1})\, z_t\right\} = 0, \tag{6}$$

with $\phi = (\theta + \omega - \omega\theta + \omega\beta\theta)$.

The vector z_t includes variables that are orthogonal to υ_{t+1}, allowing for GMM estimation. Quarterly U.S. data are used, with π_t measured by the percentage change in the GDP deflator, and real marginal costs given by the logarithm of the labour income share. Finally, their instruments include four lags of inflation, labour share, commodity-price inflation, wage inflation, the long-short interest rate spread, and the output gap (measured by a detrended log GDP).

Galí and Gertler's estimations yield values for ω, θ, and β: (0.27, 0.81, 0.89) for specification (1), and (0.49, 0.83, 0.91) for specifications (2). Where the subjective discount rate is restricted to one, the estimates are (0.24, 0.80, 1.00) and (0.52, 0.84, 1.00), respectively. The implied slopes are all positive and found to be significant judging from the IV-based asymptotic standard errors, and given the fact that the overidentifying restrictions are not rejected relying on the J-test. Accordingly, Galí and Gertler conclude that there is good empirical support for the NKPC, and, furthermore, that the forward-looking component of inflation is more important than the backward-looking component; i.e., the estimated γ_f is larger than the estimated γ_b.

Given that the severity of weak-instruments effects is now well understood in econometrics, however, it is important to ascertain that these results are not invalidated by such problems.[6] Ma (2002) uses corrected GMM inference methods developed by Stock and Wright

[5] In Galí and Gertler (1999), the orthogonality conditions are written for the case $\omega = 0$; see Galí, Gertler, and Lopez-Salido (2001) for the general case.

[6] For details on weak instruments and their effects, see Nelson and Startz 1990a,b; Buse 1992; Choi and Phillips 1992; Maddala and Jeong 1992; Angrist and Krueger 1994; McManus, Nankervis, and Savin 1994; Bound, Jaeger, and Baker 1995; Cragg and Donald 1996; Hall et al. 1996; Dufour 1997; Staiger and Stock 1997; Wang and Zivot 1998; Zivot, Startz, and Nelson 1998; Stock and Wright 2000; Dufour and Jasiak 2001; Hahn and Hausman 2003; Kleibergen 2002; Moreira 2003a,b; Stock, Wright, and Yogo 2002; Kleibergen and Zivot 2003; and Wright 2003. Several works are also cited in Dufour (2003) and Stock, Wright, and Yogo (2002).

(2000) to re-evaluate the empirical relevance of the NKPC specifications. The corrected 90 per cent confidence sets (called S-sets) that Ma calculates are very large, including all parameter values between [0, 3] for two of the structural parameters, and [0, 8] for the third. The fact that all parameter combinations derived from these value ranges are compatible with the model suggests that the parameters are weakly identified. We next reassess the NKPC model using identification-robust (or robust to weak instruments) methods.

3. Statistical Framework and Methodology

We consider two econometric specifications to assess Galí and Gertler's NKPC:

$$\pi_t = \lambda s_t + \gamma_f \pi_{t+1} + \gamma_b \pi_{t-1} + u_{t+1}, \qquad t = 1, ..., T \tag{7}$$

and

$$\pi_t = \lambda s_t + \gamma_f \tilde{\pi}_{t+1} + \gamma_b \pi_{t-1} + u_{t+1}, \qquad t = 1, ..., T \tag{8}$$

where $\tilde{\pi}_t$ is a survey measure for inflation expectations.

The two models differ in their assumptions regarding the formation of inflation expectations. In (7), expected inflation, $E_t \pi_{t+1}$, is proxied by the realized value, π_{t+1}, while in (8) it is replaced by the survey-based measure, $\tilde{\pi}_{t+1}$, of expected inflation. It is easy to see that both approaches raise error-in-variable problems and the possibility of correlation between explanatory variables and the disturbance term in the two above equations. Researchers such as Roberts (1997, 1998, 2001) have noted that the maintained specification for how expectations are formed has important implications for the empirical validity of the curve. That is, additional lags not implied by the NKPC under rational expectations may be required, even if the model is conditional on labour costs.

The parameters λ, γ_f, and γ_b, defined in equation (3), are nonlinear transformations of the "deep parameters" ω, β, and θ. The statistical details underlying our inference methodology are provided in Appendix B, where, to simplify the presentation, we adopt the following notation: y is the T-dimensional vector of observations on π_t, Y is the $T \times 2$ matrix of observations on s_t and either of π_{t+1} and $\tilde{\pi}_{t+1}$, X_1 is the vector of observations on the inflation lag π_{t-1}, X_2 is the $T \times k_2$ matrix of the instruments (we use 24 instruments; see section 4) and u is the T-dimensional vector of error terms, u_t.

The methodology we consider can be summarized as follows. To obtain a confidence set with level $(1 - \alpha)$ for the deep parameters, we invert the F-test described in Appendix B associated with the null hypothesis:

$$H_0 : \omega = \omega^0, \beta = \beta^0, \theta = \theta^0, \text{ where } \omega^0, \beta^0, \text{ and } \theta^0 \text{are known values.} \tag{9}$$

Formally, this implies collecting the values ω^0, β^0, and θ^0 that are not rejected by the test (i.e., the values for which the test is not significant at level α). Taking equation (8) as an example, the test under consideration proceeds as follows (further discussion and references are provided in Appendix B).

(i) Applying (3) and (4), obtain the values of λ, γ_f, and γ_b associated with ω^0, β^0, and θ^0; denote these transforms λ^0, γ_f^0, and γ_b^0.

(ii) Consider the regression (which we denote the AR regression, following Anderson and Rubin 1949) of

$$\left\{\pi_t - \lambda^0 s_t - \gamma_f^0 \tilde{\pi}_t - \gamma_b^0 \pi_{t-1}\right\} \text{ on } \{\pi_{t-1} \text{ and the } instruments\}. \tag{10}$$

Under the null hypothesis—specifically, (8) and (9)—the coefficients of the latter regression should be zero. Hence, testing for a zero null hypothesis on all response coefficients in (10) provides a test of (9).

(iii) Compute the standard F-test associated with the exclusion of all regressors, namely,

$$\{\pi_{t-1} \text{ and the } instruments\}$$

in the regression (10); see (23) in Appendix B. In this context, the usual classical regression framework applies, so the latter F-test can be referred to its usual F or χ^2 cut-off points.

Tests of this type were originally proposed by Anderson and Rubin (1949) for linear Gaussian simultaneous equations models. The AR approach transforms a structural equation such as (8) into the regular regression framework as in (10), for which standard finite-sample and asymptotic distributional theory applies. The required transformation is extremely simple, despite the complexity of the model under test. Indeed, the basic test we use for inference on ω^0, β^0, and θ^0 differs from a standard IV-based Wald or t-type test, in that it avoids estimating the structural equation in its (8) form, which faces identification difficulties. In contrast, the AR-regression (10) satisfies the usual classical regression assumptions (because no "endogenous" variables appear on its right-hand side). Whereas any statistical analysis of (8) requires identification constraints, these are no longer needed for inference on (10). As is shown more rigorously in Appendix B, the AR regression provides information on the structural parameters because it is linked to the statistical reduced form associated with the structural equation (8). By identification-robustness we mean that the F-test is valid whether the model is identified or not.[7]

Transforming the test problem to the AR-regression framework comes at some cost: the identification-robust F-test requires assessing—in the regression (10)—the exclusion of π_{t-1} and the 24 available instruments (25 constraints), even though only three structural parameters are being tested. Instrument abundance thus leads to degrees-of-freedom losses, with obvious consequences for test power. It is possible to characterize what an "optimal" instrument set looks like from the point of view of maximizing test power: up to a non-singular transformation, the latter (say, $\overline{Z}$) should be the mean of the endogenous explanatory variables in the model, or, which is equivalent,

$$X_2 \times \{\text{the coefficient of } X_2 \text{ in the first-stage regression, assumed known}\},$$

[7] We emphasize in Appendix B that the latter test will be exactly size correct if we can strictly condition on the regressors and, particularly, the instruments for statistical analysis; weakly exogenous regressors in our dynamic model with instruments orthogonal to the regression error terms are not in accord with the latter assumption. Nevertheless, the tests are still identification-robust. An exact test can still be devised for the NKPC model at hand, despite its dynamic econometric specification, if one is willing to consider strongly exogenous instruments.

where X_2 (as defined above) refers to the matrix of available *instruments*; see Dufour and Taamouti 2003b and Appendix B). The first-stage regression is the regression of the left-hand-side endogenous variables in (8)—marginal cost and expected inflation—on the included exogenous variable (the inflation lag) and X_2. Formally, this implies applying the above steps (i) to (iii), replacing the *instruments* by $\overline{Z}$, and whose dimension is $T \times 2$. Therefore, the optimal identification-robust F-test requires assessing, in the regression (10), the exclusion of π_{t-1} and the two optimal instruments (three constraints); recall that the number of structural parameters under test is indeed three. This provides optimal information reduction, which improves the power of the test (and thereby may tighten the confidence sets based on these tests).

In practice, however, the coefficient of X_2 in the first-stage regression (Π_2 in Appendix B) is not known, estimates of this parameter must be "plugged in," which of course leads to an "approximately optimal" procedure. As described in Dufour (2003), many procedures that aim at being identification-robust as well as improving the AR procedure from the viewpoint of power rely on different choices of $\overline{Z}$. In particular, if a constrained OLS estimator imposing the structure underlying equation (8)—denoted $\hat{\Pi}_2^0$ in Appendix B; see (25)— is used, then the associated procedure yields Kleibergen's (2002) K-test.[8] In other words, Kleibergen's (2002) K-test applies the above steps (i) to (iii), replacing the *instruments* by

$$\bar{Z}_K = X_2 \hat{\Pi}_2^0.$$

To avoid confusion, the tests based on X_2 and $\bar{Z}_K$ are denoted by AR and AR-K, respectively.

We invert these tests to get confidence sets as follows: using a grid search over the economically meaningful set of values for ω, β, and θ, we sweep relevant choices for ω^0, β^0, and θ^0.[9] For each parameter combination considered, we compute the statistics AR and AR-K and their respective p-values. The parameter vectors for which the p-values are greater than the level α constitute a confidence set with level $(1 - \alpha)$. Since every choice of ω^0, β^0, and θ^0 entails—using (3)—a choice for λ, γ_f, and γ_b, this procedure also yields conformable confidence sets for the latter parameters. These confidence sets reflect the structure, and are obtained without further computations, although λ, γ_f, and γ_b are transformations of the deep parameters. Therein lies a significant advantage in using our approach as an alternative to standard nonlinear Wald-based techniques.

In summary, two points are worth emphasizing. First, if the confidence set we obtain by inverting any of these AR-type tests is empty (i.e., no economically acceptable value of the model's deep parameters is upheld by the data), then we can infer that the model is rejected at the chosen significance level. Our procedure can therefore be seen as an identification-robust alternative to the standard GMM-based J-test. In the same vein, utterly uninformative (too-wide) confidence sets also allow the model fit to be assessed, since unbounded confidence sets do occur under identification difficulties (see the discussion in Dufour 2003). Our procedure

[8] To correct for plug-in estimation effects (i.e., to estimate Π_2), Dufour and Jasiak (2001) and Dufour and Taamouti (2003a,b) recommend split sample estimation techniques, where the first subsample is used to calculate Π_2 and the second subsample to run the optimal AR test with the latter estimate. The results obtained by applying these versions of the tests are available from the authors upon request.

[9] We allow the range (0, 1) as the admissible space for each of ω, θ, and β. The values are varied with increments of 0.03 for ω and θ, and by 0.01 for β. The increment of 0.03 was chosen for the first two parameters (rather than 0.01), to minimize the computational burden.

(which performs, for practical purposes, the same specification checks as a J-type test) has a clear "built-in" advantage relative to GMM-based t-type confidence intervals, backed by a non-significant J-test.[10]

Our procedure offers a second important advantage not shared by the latter standard approach. So far, we have considered the estimation and test problem given a specific significance (or confidence) level, α. Alternatively, the p-value associated with the above defined tests, which provides a formal specification check, can be used to assess the empirical fit of the model. In other words, the values (uniqueness is not granted) of ω^0, β^0, and θ^0 that lead to the largest p-value formally yield the set of "least rejected" models; i.e., models that are most compatible with the data.[11] In practice, analyzing the economic information content of these least rejected models (associated with the least rejected "deep parameter" combinations) provides decisive and very useful goodness-of-fit checks.

4. Empirical Results

We applied the above-defined inference methods to the hybrid NKPC models in (7) and (8) for both U.S. and Canadian data. One difference between our specifications and those of Galí and Gertler is that we use a real-time output-gap measure in the set of instruments, instead of a gap detrended using the full sample. The latter measure is likely not an appropriate instrument since, when the full sample is used, lagged values of the gap are, by construction, related to future information. To avoid this, we proceed iteratively: to obtain the value of the gap at time t, we detrend GDP with data ending in t. The sample is then extended by one observation and the trend is re-estimated. This is used to detrend GDP and yields a value for the gap at time $t+1$. This process is repeated until the end of the sample. In this fashion, our gap measures at time t do not use information beyond that period and it can therefore be used as valid instruments. A quadratic trend is used for this purpose.[12]

Regarding survey expectations, the Federal Reserve Bank of Philadelphia publishes quarterly mean forecasts of the next quarter's U.S. GDP implicit price deflator. We first-difference this series to obtain our inflation-expectations series for the United States.[13] In the case of Canada, the survey-based inflation expectations series were obtained from the Conference Board of Canada survey; further details on the Canadian data are provided in the Appendix. For the remaining variables, other than the output gap, we use the Galí and Gertler data and instrument set for the United States, and the corresponding variables in the case of Canada. Because of the expectations variables in the data set, our samples start in 1970Q1.

We first apply the AR test to the U.S. data and to equation (7), to assess the Galí and Gertler (1999) reported estimates. Specifically, we test whether ω, θ, and β are (0.27, 0.81, 0.89) or (0.49, 0.83, 0.91), which correspond to those authors' estimates for their orthogonality

[10]Indeed, if the AR confidence set with level $(1-\alpha)$ is empty, then the usual LIML overidentification test statistic will exceed a specific bounds-based identification-robust α-level critical point; i.e., the associated overidentification test is conclusively significant at level α.

[11]This method underlies the principles of Hodges-Lehmann estimators; see Hodges and Lehmann 1963, 1983). Least rejected values may thus be interpreted as "point estimates."

[12]We repeated our estimations using a cubically detrended real-time gap measure, as well as the Christiano-Fitzgerald one-sided band-pass filter, and obtained qualitatively similar results.

[13]Source: <http://www.phil.frb.org/econ/spf/index.html>.

Table 1: Anderson-Rubin Tests with Rational Expectations

Test type		Unrestricted model			
		Max p-value	Deep parameters (ω, θ, β)	Reduced-form parameters (λ, γ_f, γ_b)	Freq.
AR	U.S.	0.2771	(0.40, 0.64, 0.96)	(0.08, 0.60, 0.39)	2.78
	Canada	-	-	-	-
AR-K	U.S.	0.9993	(0.40, 0.61, 0.98)	(0.09, 0.59, 0.40)	2.56
	Canada	0.9990	(0.01, 0.37, 0.21)	(1.53, 0.21, 0.03)	1.59
		$\beta = 0.99$			
		Max p-value	Deep parameters (ω, θ, β)	Reduced-form parameters (λ, γ_f, γ_b)	Freq.
AR	U.S.	0.2765	(0.37, 0.64, 0.99)	(0.08, 0.63, 0.37)	2.78
	Canada	-	-	-	-
AR-K	U.S.	0.9987	(0.37, 0.64, 0.99)	(0.08, 0.63, 0.37)	2.78
	Canada	0.2900	(0.01, 0.10, 0.99)	(7.30, 0.91, 0.09)	1.11

Notes: AR is the Anderson-Rubin test and AR-K refers to the Kleibergen approximate optimal instruments test. Freq. is the average frequency of price adjustment, measured in quarters.

specifications (1) and (2), respectively. We find all tests to be significant at conventional levels, so that Galí and Gertler's estimates are rejected. We then ask whether, for the same instrument set, there are any parameter combinations for which the hybrid NKPC is not rejected. Interestingly, we find some dramatically different results depending on whether model (7) or (8) is used.

For the U.S. rational-expectations solution, we find a bounded but fairly large confidence set. This means that a multitude of different parameter combinations are compatible with the econometric model tested, although the set is much smaller than the S-sets constructed by Ma.[14] However, for the model using survey expectations, the confidence set is empty at the 95 per cent level. Thus, not a single parameter value combination is compatible with this particular econometric model, which implies that, with survey expectations, the model is not identified. Regarding the Canadian data, we find that the outcomes are reversed. Thus, it is the model with rational expectations that generates the empty confidence set, while the specification using survey data yields the non-empty one. The latter is so small that there are only some parameter value combinations for which the model is statistically valid.

Along with the identification-robust confidence sets, one of the great advantages of using the Anderson-Rubin method is that it yields the parameter combination that is least rejected, or, alternatively, that has the highest p-value. Formally, as explained in the previous section, this point estimate corresponds to the so-called Hodges-Lehmann estimate and it can be compared with point estimates obtained using more conventional estimation methods (such as GMM). We report this estimate for the United States and Canada in the upper panels of Tables 1 and 2, respectively. From here, we can see that, under rational expectations, the values of the deep parameters that correspond to the maximal p-value for the United States are given by the set (0.40, 0.64, 0.96) for (ω, θ, β), respectively. These translate into a value of 0.6 for the coefficient of the forward-looking component of inflation, γ_f, and 0.39 for the coefficient of the backward-looking component, γ_b. Furthermore, the coefficient on the marginal cost variable is 0.08, and the average frequency of price adjustment is 2.78 quarters.

Based on the Hodges-Lehmann estimates, the findings provide support for the optimization-based Phillips curve, and the notion that the forward-looking component of the U.S. inflation process is more important than its backward-looking component. In addition, the estimate for the average frequency of price adjustment is fairly close to the value of 1.8 obtained based on micro data (see for example Bils and Klenow (2004).[15] On the other hand, the graphs in the second panel of Figure 1 add a qualifier to the above statement.

In Figure 1, the graph in the bottom panel, on the left, depicts the 95 per cent (solid line, p-value= 0.05) and 90 per cent (dashed line, p-value= 0.10) confidence sets based on the AR test, and for the case where the subjective discount parameter is constrained to lie between 0.95 and 0.99. An "X" marks the spot corresponding to the highest p-value obtained (0.2797). Immediately, three features can be observed: (i) the set of parameter values that the test does not reject at the 5 and 10 per cent levels are fairly large, (ii) within these sets, there is more than one ω value that corresponds to a given θ, and vica versa, and (iii) the parameter combination that yields the highest p-value is very close to points that have a p-value of 0.10 only. In other words, even when β is constrained quite tightly, the

[14]There is a slight difference between our two instrument sets: Ma's set includes a constant and has no fourth lag for any of the variables in levels.

[15]Galí and Gertler report average price adjustment frequencies of about 4 to 6 quarters.

Table 2: Anderson-Rubin Tests with Survey Expectations

Test type		Unrestricted model			
		Max p-value	Deep parameters (ω, θ, β)	Reduced-form parameters (λ, γ_f, γ_b)	Freq.
AR	U.S.	-	-	-	-
	Canada	0.1009	(0.01, 0.97, 0.89)	(0.00, 0.88, 0.01)	33.33
AR-K	U.S.	0.9983	(0.01, 0.61, 0.64)	(0.38, 0.63, 0.02)	2.56
	Canada	0.0890	(0.01, 0.97, 0.90)	(0.00, 0.89, 0.01)	33.33
		$\beta = 0.99$			
		Max p-value	Deep parameters (ω, θ, β)	Reduced-form parameters (λ, γ_f, γ_b)	Freq.
AR	U.S.	-	-	-	-
	Canada	0.0562	(0.01, 0.97, 0.99)	(0.00, 0.98, 0.01)	33.33
AR-K	U.S.	0.6057	(0.52, 0.22, 0.99)	(0.40, 0.29, 0.70)	1.28
	Canada	-	-	-	-

Notes: AR is the Anderson-Rubin test and AR-K refers to the Kleibergen approximate optimal instruments test. Freq. is the average frequency of price adjustment, measured in quarters.

uncertainty regarding the estimated values of the other parameters is relatively high. This is seen more easily in the adjacent graph, which depicts the values corresponding to the 95 per cent confidence set in the γ_f and γ_b space. Notice, in particular, that a value of 0.60 for the backward-looking component of inflation, and 0.37 for the forward-looking component is as likely to be obtained as a value of 0.90 and 0.10 for the forward- and backward-looking components, respectively.

Turning to the case of the AR test applied to Canadian data, recall that the model with rational expectations is not compatible with the data, but that the model with survey expectations does yield a non-empty set. The results corresponding to the highest p-value for the latter are reported in Table 2. In this case, the maximal p-value is 0.1009, while the deep parameters are (0.01, 0.97, 0.89) for (ω, θ, β), respectively. Based on the fact that the proportion of firms that follow a rule-of-thumb is practically zero ($\omega = 0.01$), we would conclude that a purely forward-looking model is applicable to Canada. However, the reduced-form parameters and the average frequency of price adjustment indicate that the model is economically not viable. This is the case even if β is constrained to 0.99 in the estimation.[16]

Results based on Kleibergen's statistic are also reported in Tables 1-2. As for the AR tests, two sets of outcomes are tabulated: the parameter values that yield the highest test p-value for the unrestricted model appear in the upper panel of Tables 1 and 2, while in the lower panel shows the correponding elements but with β constrained to 0.99.

Let us first examine the results for the United States with the rational-expectations model. When $\bar{Z}_K$ is used as the optimal instrument set, the model is least rejected for the parameter combination (0.40, 0.61, 0.98), and the p-value is 0.9993. These values are extremely close to those reported for the correponding restricted-estimation (with β constrained to 0.99) case, and to those reported for the AR tests.

With the model based on survey expectations (Table 2), although the AR test yields an empty confidence set for the United States, the AR-K test yields a least-rejected parameter combination that suggests strongly forward-looking behaviour ($\gamma_f = 0.63$, $\gamma_b = 0.02$). In addition, when the subjective discount rate is constrained to 0.99, the AR-K test points to a much more important backward-looking component for inflation.

Our findings are somewhat similar with Canadian data. Thus, although the AR-K test yields similar outcomes to those of the AR test for the unrestricted model with survey expectations, with rational expectations, the AR-K test yields parameter values that suggest a more important backward-looking role in inflation. In addition, the estimate for the average frequency of price adjustment is 1.6, very much in line with micro data (as in Bils and Klenow 2004). These results are nevertheless difficult to reconcile with the value for ω, which is essentially zero. In addition, once the subjective discount rate parameter is constrained to 0.99, the conclusions for the rational-expectations specification from the AR-K test point to a much more important forward-looking component of inflation ($\gamma_f = 0.91$, $\gamma_b = 0.09$). The unusual feature in this case is the value of the coefficient on the marginal cost variable, λ, which stands at 7.30.

Figures 1 and 2 show U.S. results for the AR-K test for the case where β is constrained

[16] For this reason, and because all of the admissible ω values in the AR-based confidence sets equal 0.01, no figures are provided for Canada.

Figure 1: AR and AR-K Tests (U.S., Rational Expectations)

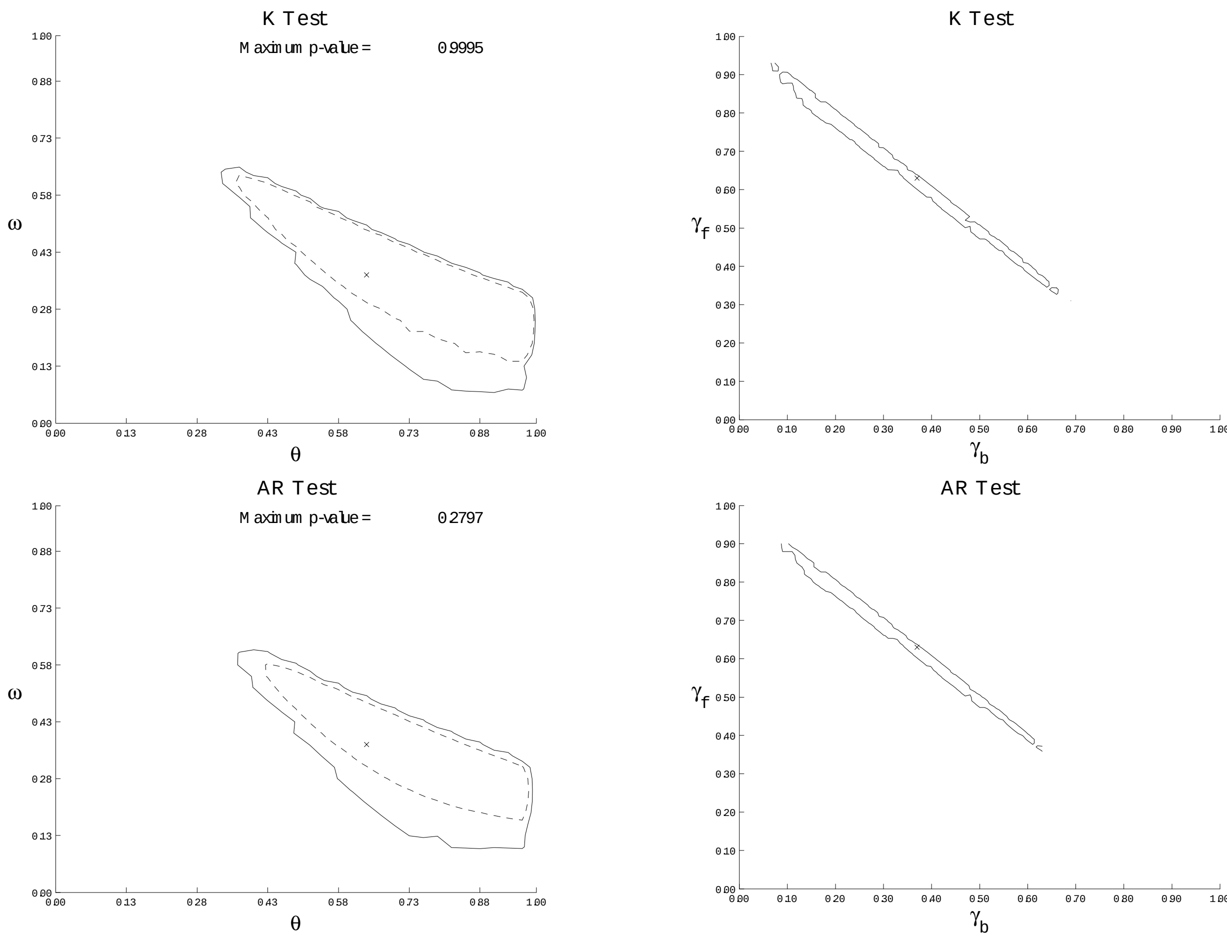
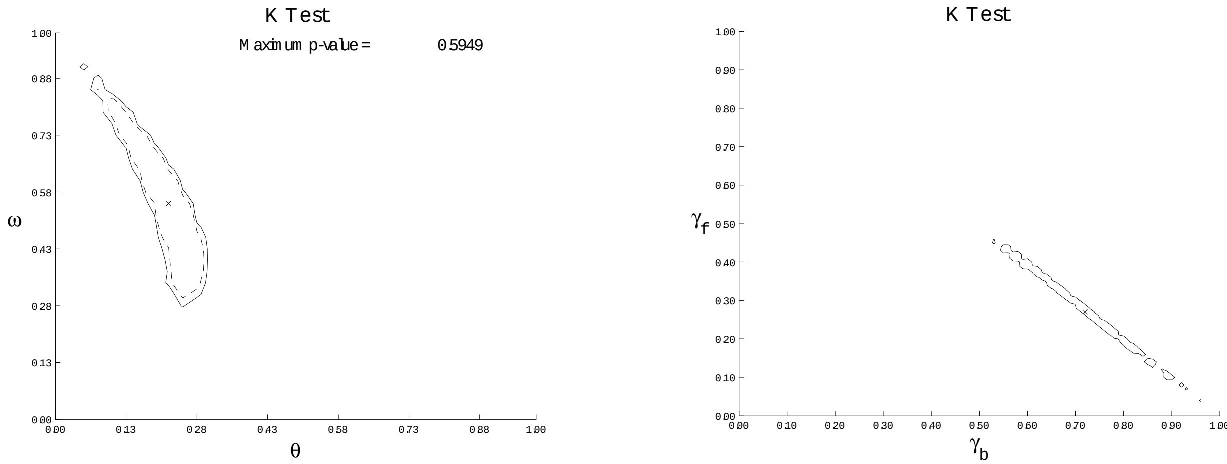


Figure 2: AR-K Tests (U.S., Survey Expectations)



to fall between 0.95 and 0.99. With rational expectations (Figure 1), interestingly, the confidence set based on inverting the AR-K test is larger than that based on AR, but the results are in line with each other, in the sense that the 95 per cent confidence sets are more skewed towards higher γ_f than γ_b. Turning to Figure 2, we find that the AR-K test produces strong support for a larger backward-looking component to inflation.

Taken collectively, our results point to problems of weak identification in these models. Nevertheless, we find that there is some support for the hybrid NKPC for the United States, whereas the model is not suited to Canada.

5. Conclusion

In this paper, we used finite-sample methods to test the empirical relevance of Galí and Gertler's (1999) NKPC equations, using AR tests as well as Kleibergen's more parsimonious procedure, for both U.S. and Canadian data. Two variants of the model were studied: one based on a rational-expectations assumption, and a modification to the latter that uses survey data on inflation expectations. In the U.S. case, Galí and Gertler's (1999) original data set was used except for the output-gap measure and survey expectations, where applicable.

First, we found some evidence of identification difficulty. Nevertheless, the maximal p-value arguments selected those parameter values for which the model is least rejected—a very useful feature of our proposed identification-robust techniques. Second, we found support for Galí and Gertler's hybrid NKPC specification with rational expectations for the United States. Third, neither model was found to be well suited to describe inflation dynamics in Canada. Fourth, we found that, for the cases where the Anderson-Rubin test yields an empty confidence set, the AR-K procedure leads to conflicting results for the restricted and unrestricted models.

These results underscore the need for employing identification-robust inference in the estimation of expectations-based dynamic macroeconomic relations.

References

Anderson, T.W. and H. 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, J.D. and A.B. Krueger. 1994. "Split Sample Instrumental Variables." NBER Working Paper No. 150. Cambridge, MA.

Bils, M. and P. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112: 947–85.

Bound, J., D.A. Jaeger, and R.M. Baker. 1995. "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90: 443–50.

Buse, A. 1992. "The Bias of Instrumental Variables Estimators." *Econometrica* 60: 173–80.

Choi, I. and P.C.B. Phillips. 1992. "Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations." *Journal of Econometrics* 51: 113–50.

Cragg, J.G. and S.G. Donald. 1996. "Testing Overidentifying Restrictions in Unidentified Models." Department of Economics, University of British Columbia and Boston University. Discussion Paper No. 96-20.

Dufour, J.M. 1997. "Some Impossibility Theorems in Econometrics, with Applications to Structural and Dynamic Models." *Econometrica* 65: 1365–89.

———. 2003. "Identification, Weak Instruments and Statistical Inference in Econometrics." *Canadian Journal of Economics* 36(4): 767–808.

Dufour, J.M. and J. Jasiak. 2001. "Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors." *International Economic Review* 42: 815–43.

Dufour, J.M. and L. Khalaf. 2002. "Exact Tests for Contemporaneous Correlation of Disturbances in Seemingly Unrelated Regressions." *Journal of Econometrics* 106(1): 143–70.

———. 2003a. "Finite Sample Tests in Seemingly Unrelated Regressions." In *Computer-Aided Econometrics*, edited by D.E.A. Giles, 11–35. New York: Marcel Dekker.

———. 2003b. "Simulation-Based Finite-Sample Inference in Simultaneous Equations." Technical report. CIRANO and CIREQ, Université de Montréal.

Dufour, J.M. and M. Taamouti. 2003a. *On Methods for Selecting Instruments*. Technical report, C.R.D.E., Université de Montréal.

———. 2003b. *Point-Optimal Instruments and Generalized Anderson-Rubin Procedures for Nonlinear Models*. Technical report, C.R.D.E., Université de Montréal.

———. 2004. *Further Results on Projection-Based Inference in IV Regressions with Weak, Collinear or Missing Instruments*. Technical report, Département de sciences économiques, Université de Montréal.

———. 2005. "Projection-Based Statistical Inference in Linear Structural Models with Possibly Weak Instruments." *Econometrica* Forthcoming.

Eichenbaum, M. and J. Fisher. 2004. "Evaluating the Calvo Model of Sticky Prices." Working Paper. Federal Reserve Bank of Chicago.

Fuhrer, J. 1997. "The (Un)importance of Forward-Looking Behaviour in Price Specifications." *Journal of Money, Credit and Banking* 29: 338–50.

Fuhrer, J. and G. Moore. 1995. "Inflation Persistence." *Quarterly Journal of Economics* 110: 127–59.

Fuhrer, J.C. and G.P. Olivei. 2004. *Estimating Forward Looking Euler Equations with GMM Estimators: An Optimal Instruments Approach*. Technical report, Federal Reserve Bank of Chicago.

Gagnon, E. and H. Khan. 2005. "New Phillips Curve under Alternative Production Technologies for Canada, the U.S., and the Euro Area." *European Economic Review* 49: 1571–1602.

Galí, J. and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44: 195–222.

Galí, J., M. Gertler, and J.D. Lopez-Salido. 2001. "European Inflation Dynamics." *European Economic Review* 45: 1237–70.

Hahn, J. and J. Hausman. 2003. *IV Estimation with Valid and Invalid Instruments*. Technical report, Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hall, A.R., G.D. Rudebusch, and D.W. Wilcox. 1996. "Judging Instrument Relevance in Instrumental Variables Estimation." *International Economic Review* 37: 283–298.

Hodges, J.L., Jr. and E.L. Lehmann. 1963. "Estimates of Location Based on Rank Tests." *Annals of Mathematical Statistics* 34: 598–611.

———. 1983. "Hodges-Lehmann Estimators." In *Encyclopedia of Statistical Sciences, Volume 3*, edited by N.L. Johnson, S. Kotz, and C. Read, 642–45. New York: John Wiley & Sons.

Khalaf, L. and M. Kichian. 2004. "Estimating New Keynesian Phillips Curves Using Exact Methods." Bank of Canada Working Paper 2004-11.

———. 2005. "Exact Tests of the Stability of the Phillips Curve: The Canadian Case." *Computational Statistics and Data Analysis* 49.

Kleibergen, F. 2002. "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression." *Econometrica* 70(5): 1781–1803.

Kleibergen, F. and E. Zivot. 2003. "Bayesian and Classical Approaches to Instrumental Variable Regression." *Journal of Econometrics* 114(1): 29–72.

Linde, J. 2001. *Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach.* Technical report, Sveriges Riksbank No. 129., Stockholm, Sweden.

Ma, A. 2002. "GMM Estimation of the New Phillips Curve." *Economic Letters* 76: 411–17.

Maddala, G.S. and J. Jeong. 1992. "On the Exact Small Sample Distribution of the Instrumental Variable Estimator." *Econometrica* 60: 181–83.

McManus, D.A., J.C. Nankervis, and N.E. Savin. 1994. "Multiple Optima and Asymptotic Approximations in the Partial Adjustment Model." *Journal of Econometrics* 62: 91–128.

Moreira, M.J. 2003a. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71(4): 1027–48.

———. 2003b. *A General Theory of Hypothesis Testing in the Simultaneous Equations Model.* Technical report, Department of Economics, Harvard University, Cambridge, Massachusetts.

Nason, J.M. and G.W. Smith. 2003. *Identifying the New Keynesian Phillips Curve.* Technical report, Univeristy of British Columbia and Queen's University.

Nelson, C.R. and R. Startz. 1990a. "The Distribution of the Instrumental Variable Estimator and its t-ratio when the Instrument is a Poor One." *Journal of Business* 63: 125–40.

———. 1990b. "Some Further Results on the Exact Small Properties of the Instrumental Variable Estimator." *Econometrica* 58: 967–76.

Roberts, J. 1997. "Is Inflation Sticky?" *Journal of Monetary Economics* 39: 173–96.

———. 1998. *Inflation Expectations and the Transmission of Monetary Policy.* Technical report, Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 1998-43, Washington, D.C.

———. 2001. *How Well Does the New Keynesian Sticky-Price Model Fit the Data?* Technical report, Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 2001-13, Washington, D.C.

Staiger, D. and J.H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3): 557–86.

Stock, J.H. and J.H. Wright. 2000. "GMM with Weak Identification." *Econometrica* 68: 1097–1126.

Stock, J.H., J.H. Wright, and M. Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20(4): 518–29.

Wang, J. and E. Zivot. 1998. "Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments." *Econometrica* 66(6): 1389–1404.

Woodford, M. 2003. *Interest and Prices*. Princeton, New Jersey: Princeton University Press.

Wright, J.H. 2003. "Detecting Lack of Identification in GMM." *Econometric Theory* 19(2): 322–30.

Zivot, E., R. Startz, and C.R. Nelson. 1998. "Valid Confidence Intervals and Inference in the Presence of Weak Instruments." *International Economic Review* 39: 1119–44.

Appendix A: Data Description for Canada

The inflation-expectations series is obtained from the Conference Board of Canada survey. Each quarter, participants are asked to forecast the annual average (GDP-deflator) inflation rate for the current year. Let us denote $\tilde{\pi}^a_1$, $\tilde{\pi}^a_2$, $\tilde{\pi}^a_3$, and $\tilde{\pi}^a_4$ the annual average inflation forecasts made in quarters 1, 2, 3, and 4 of a given year, respectively. Clearly, forecasts that are made in the second, third, and fourth quarters are likely to integrate realized (and observed) inflation in quarters 1, 1 and 2, and 1, 2, and 3, respectively.

To obtain a "pure" quarterly expectations series, we proceed as follows: First, denote the forecasted quarterly inflation rate in quarters 1 to 4 as $\tilde{\pi}^q_1$, $\tilde{\pi}^q_2$, $\tilde{\pi}^q_3$, and $\tilde{\pi}^q_4$, respectively. Similarly, let π^q_1, π^q_2, and π^q_3 be the realized quarterly inflation rates in quarters 1, 2, and 3, respectively. Then, the forecasted quarterly inflation rates are calculated as follows:

$$\begin{aligned}
\tilde{\pi}^q_1 &= \tilde{\pi}^a_1/4, \\
\tilde{\pi}^q_2 &= (\tilde{\pi}^a_2 - \pi^q_1)/3, \\
\tilde{\pi}^q_3 &= (\tilde{\pi}^a_3 - \pi^q_1 - \pi^q_2)/2, \\
\tilde{\pi}^q_4 &= (\tilde{\pi}^a_4 - \pi^q_1 - \pi^q_2 - \pi^q_3).
\end{aligned}$$

The remaining data are quarterly time series from Statistics Canada's database. Any monthly data are converted to quarterly frequency.

Output gap is the deviation of real GDP ($y_t = \ln Y_t$) from its steady state, approximated by a quadratic trend: $\hat{y} = 100(y_t - \bar{y}_t)$, where $Y_t = I56001 - I56013 - I56018$.

Price inflation is the quarterly growth rate of the total GDP deflator:
$\pi_t = 100(\ln P_t - \ln P_{t-1})$ and $P_t = D15612$.

Wage inflation is the quarterly growth rate of compensation of employees:
$w_t = 100(\ln W_t - \ln W_{t-1})$, where $W_t = D17023/N_t$.
$N_t = LFSA201$ for 1970Q1–1975Q4 and $N_t = D980595$ for 1976Q1–2000Q4.

Labour income share is the ratio of total compensation and nominal GDP: $ls_t = \ln S_t$, and $s_t = 100(ls_t - s)$, the labour income share in deviation from its steady state, where $s = \ln S, S = \sum_t^T \ln(S_t)/T$, and $S_t = (D17023 - D17001)/(D15612 * Y_t)$.

Average real marginal costs for Cobb-Douglas Function: $rmc^{avg}_t = s_t$.

Appendix B: The AR Test and Related Procedures

Consider the following structural equation:

$$y = Y\delta + X_1\kappa + u, \tag{11}$$

where y is a $T \times 1$ dependent variable, Y is a $T \times m$ matrix of endogenous variables, X_1 is a $T \times k_1$ matrix of exogenous variables, and u is an error term that satisfies standard regularity conditions typical of IV regressions (Dufour and Jasiak 2001). In our context (see section 3), y is the T-dimensional vector of observations on π_t, Y is the $T \times 2$ matrix of observations on s_t and $\tilde{\pi}_{t+1}$ (or π_{t+1}, depending on the context), X_1 is the vector of observations on the inflation lag π_{t-1}, X_2 is the $T \times k_2$ matrix of the instruments, and u is the T-dimensional vector of error terms, u_t.

Suppose that the reduced form associated with the right-hand-side endogenous regressors is

$$Y = X_1\Pi_1 + X_2\Pi_2 + V, \tag{12}$$

where V is a $T \times m$ matrix of error terms assumed to be cross-correlated and correlated with u, and X_2 is the matrix of available instruments.[17] In this case, the reduced form associated with (11) is

$$y = X_1 p_1 + X_2 p_2 + u + V\delta, \tag{13}$$

$$p_1 = \Pi_1\delta + \kappa, \qquad p_2 = \Pi_2\delta. \tag{14}$$

Identification constraints follow from (14) and amount to the rank condition

$$rank(\Pi_2) = m. \tag{15}$$

Consider hypotheses of the form

$$H_0 : \delta = \delta^0. \tag{16}$$

In this case, the model transformed as follows

$$y - Y\delta^0 = Y(\delta - \delta^0) + X_1\kappa + u,$$

has the reduced form

$$y - Y\delta^0 = X_1[\Pi_1\left(\delta - \delta^0\right) + \kappa] + X_2[\Pi_2\left(\delta - \delta^0\right)] + u + V\left(\delta - \delta^0\right). \tag{17}$$

For this reason, the AR test assesses the exclusion of X_2 (of size $T \times k_2$) in the regression of $y - Y\delta^0$ on X_1 and X_2, which can be conducted using the standard F-test. Let $X = (X_1, X_2)$, and define

$$M = I - X(X'X)^{-1}X', \qquad M_1 = I - X_1(X_1'X_1)^{-1}X_1'.$$

[17] We stress that, in Dufour and Taamouti (2003a,b,c) and Dufour (2003): (i) linearity of the latter reduced form is strictly not necessary, and (ii) further exogenous regressors ("excluded" instruments) may enter into the equation in addition to the instrument set. To present the test in its simplest form, we maintain the standard linear form (12) and refer the reader to later references for a discussion of the more general setting. Note that the assumptions regarding the reduced form for Y do not affect the actual implementation of the test, so our simplified presentation does not lack generality for practical purposes.

The statistic then takes the form

$$AR\left(\delta^0\right) = \frac{\left(y - Y\delta^0\right)'\left(M_1 - M\right)\left(y - Y\delta^0\right)/k_2}{\left(y - Y\delta^0\right)' M \left(y - Y\delta^0\right)/\left(T - k_1 - k_2\right)}. \tag{18}$$

Under the null hypothesis, and imposing strong exogeneity and identically, independently distributed normal errors,

$$AR\left(\delta^0\right) \sim F - (k_2, T - k_1 - k_2). \tag{19}$$

Following the usual classical regression analysis, the latter strong hypotheses on the error terms can be relaxed so that, under standard regularity conditions,

$$\left(k_2 \times AR\left(\delta^0\right)\right) \overset{asy}{\sim} \chi^2\left(k_2\right). \tag{20}$$

It is important to emphasize that identification constraints do not intervene here (exactly or asymptotically). In other words (19) or (20) hold whether (15) is verified or not; this is what "identification-robustness" usually means. The test can be readily extended to accommodate additional constraints on the coefficients of (the full vector or any subset of) the X_1 variables. For example, the hypothesis

$$H_0 : \delta = \delta^0,\ \kappa = \kappa^0, \tag{21}$$

can be assessed in the context of the transformed regression

$$\begin{aligned} y - Y\delta^0 - X_1\kappa^0 \quad &= \quad X_1[\Pi_1\left(\delta - \delta^0\right) + \left(\kappa - \kappa^0\right)\} \\ &\quad\ +X_2[\Pi_2\left(\delta - \delta^0\right)] + u + V\left(\delta - \delta^0\right), \end{aligned} \tag{22}$$

which leads to the following F-statistic:

$$AR\left(\delta^0,\, \kappa^0\right) = \frac{\left(y - Y\delta^0 - X_1\kappa^0\right)'\left(I - M\right)\left(y - Y\delta^0 - X_1\kappa^0\right)/(k_1 + k_2)}{\left(y - Y\delta^0 - X_1\kappa^0\right)' M \left(y - Y\delta^0 - X_1\kappa^0\right)/\left(T - k_1 - k_2\right)}. \tag{23}$$

While the test in its original form was derived for the case where the first-stage regression is linear, we re-emphasize that it is in fact robust to: (i) the specification of the model for Y, and (ii) excluded instruments; in other words, the test is valid regardless of whether the first-stage regression is linear, and whether the matrix X_2 includes all available instruments. As argued in Dufour (2003), since one is never sure that all instruments have been accounted for, the latter property is quite important. Most importantly, this test 9and several variants discussed in Dufour 2003) is the only truly pivotal statistic whose properties in finite samples are robust to the quality of instruments.

Note that exactness strictly requires that we can condition on X (i.e., we can take X as fixed for statistical analysis). This holds particularly for the instruments. In the presence of weakly exogenous regressors, the test remains identification-robust. The intuition underlying this result is the following: conducting the test via the Anderson-Rubin regressions (17)–(22), *which constitute statistical reduced forms*, easily transforms the test problem from the IV-regression (*which requires* (15)) to the classical linear regression statistical framework (*which does not require* (15)). This provides an attractive solution to identification difficulties, a property not shared by IV-based Wald statistics nor GMM-based J tests.

Despite the latter desirable statistical properties, the test as presented above provides no guidance for practitioners regarding the choice of instruments. In addition, simulation studies reported in the above-cited references show that the power of AR-type tests may be affected by the number of instruments. To see this, consider the case of (11)–(16): here, the AR test requires assessing (in the regression of $y - Y\delta^0$ on X_1 and X_2) the exclusion of the $T \times k_2$ variables in X_2, even though the number of structural parameters under test is m (the structural parameter under test δ is $m \times 1$). On recalling that identification implies $k_2 \geq m$, we see that overidentification (or, alternatively, the availability of more instruments) leads to degrees-of-freedom losses with obvious implications for power. To circumvent this problem, an optimal instrument (in the sense of *point-optimal* power) is given by

$$\bar{Z} = X_2\Pi_2,$$

where Π_2 is the coefficient of X_2 in the first-stage regression; i.e., the regression of Y on X_1 and X_2 (Dufour and Taamouti 2003b). Formally, this implies applying (19) or (23), replacing X_2 by $\overline{Z}$ (observe that X_2 intervenes in these statistics via $M = I - X(X'X)^{-1}X'$, where $X = (X_1,\, X_2)$).

Clearly, the latter optimal instrument involves information reduction, because the associated AR-test amounts to testing for the exclusion of the $T \times m$ variables in $\overline{Z}$, which preserves available degrees-of-freedom even if the model is highly overidentified. In other words, the optimal test can reflect the informational content of all available instruments with no statistical costs.

Unfortunately, Π_2 is unknown, and so the approximate optimal instruments need to be estimated, with obvious implications for feasibility and exactness. Dufour (2003) shows that if the OLS estimator,

$$\hat{\Pi}_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y, \tag{24}$$

of Π_2 in the unrestricted reduced-form multivariate regression (12) is used in the construction of $\overline{Z}$, then the associated AR-test coincides with the LM procedure defined by Wang and Zivot (1998). In addition, the K-statistic of Kleibergen (2002) may be interpreted as based on an approximation of the optimal instrument (see Dufour and Khalaf 2003b). In this case, Π_2 is replaced by its constrained reduced-form OLS estimates imposing the structural identification condition (15):

$$\hat{\Pi}_2^0 = \hat{\Pi}_2 - (X_2'M_1X_2)^{-1}X_2'M_1\left[y - Y\delta^0\right]\frac{\left[y - Y\delta^0\right]'MY}{\left[y - Y\delta^0\right]'M\left[y - Y\delta^0\right]}. \tag{25}$$

Wang and Zivot (1998) show that the distribution of the LM statistic is bounded by the $\chi^2(k_2)$ distribution; Kleibergen (2002) shows that a $\chi^2(m)$ cut-off point is asymptotically identification-robust for the K-statistic. To obtain an $F(m, .)$ or $\chi^2(m)$ cut-off point for both statistics correcting for plug-in effects, split sample methods (where the first subsample is used to estimate Π_2 and the second is used to run the optimal AR test based on the latter estimate) may also be exploited (see Dufour and Jasiak 2001; Dufour and Taamouti 2003b).

Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue*.**

2005

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :

Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
E-mail: publications@bankofcanada.ca
Web site: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca