

Maximum Likelihood Estimation of Dynamic Stochastic Theories with an Application to New Keynesian Pricing*

André Kurmann[†]

Université du Québec à Montréal
and CIRPÉE

September 28, 2004

Abstract

This paper proposes a novel Maximum Likelihood (ML) strategy to estimate Euler equations implied by dynamic stochastic theories. The strategy exploits rational expectations cross-equation restrictions, but circumvents the problem of multiple solutions that arises in Sargent's (1979) original work by imposing the restrictions on the forcing variable rather than the endogenous variable of the Euler equation. The paper then contrasts the proposed strategy to an alternative, widely employed method that avoids the multiplicity problem by constraining the ML estimates to yield a unique stable solution. I argue that imposing such a uniqueness condition makes little economic sense and can lead to severe misspecification. To illustrate this point, I estimate Gali and Gertler's (1999) hybrid New Keynesian Phillips Curve using labor income share as the measure of real marginal cost. My ML estimates indicate that forward-looking behavior is predominant and that the model provides a good approximation of U.S. inflation dynamics. By contrast, if the same estimates are constrained to yield a unique stable solution, forward-looking behavior becomes much less important and the model as a whole is rejected.

JEL classification: C13, E31, E32

Keywords: Maximum Likelihood, Rational Expectations, New Keynesian Phillips Curve, Inflation, Real marginal cost.

*The paper has greatly benefited from conversations with Jeff Fuhrer, Alain Guay, Jinill Kim, Bob King and Chris Otrok. The author gratefully acknowledges financial support from the Bankard Fund of the University of Virginia, the Swiss National Science Foundation (bourse de jeune chercheur), and the PAFARC research fund at UQAM.

[†]*Contact address:* André Kurmann, Université du Québec à Montréal, Département des sciences économiques, P.O. Box 8888, Downtown station, Montréal (QC) H3C 3P8, Canada. *Email:* kurmann.andre@uqam.ca.

1 Introduction

A wide variety of dynamic stochastic theories give rise to linear(ized) Euler equations between an endogenous variable y_t and a forcing variable x_t of the form

$$(1) \quad y_t = aE_t y_{t+1} + bx_t + cy_{t-1} + u_t,$$

where a , b and c are parameters implied by the theory; E_t is the expectations operator; and u_t is an unobserved error term that captures deviations from the theory. The present paper (i) exposes pitfalls of commonly used Maximum Likelihood (ML) techniques to estimate such Euler equations; (ii) proposes a novel strategy that circumvents these pitfalls; and (iii) illustrates the different points with an application to a popular New Keynesian pricing model.

The starting point of the paper is Sargent's (1979) ML estimator of the expectations theory of interest rates. Translated to the Euler equation in (1), Sargent's approach consists of proxying market expectations with a vector-autoregressive (VAR) process in x_t and y_t , and deriving the cross-equation restrictions that the theory imposes on this VAR under rational expectations. The restrictions constrain the VAR coefficients of the endogenous variable y_t (i.e. its rational expectations solution) to be a function of the parameters of the Euler equation and the remaining VAR coefficients. These unrestricted parameters and coefficients are then estimated via ML.

While straightforward in theory, implementing Sargent's approach in reality is far from trivial. Specifically, I uncover that the constrained coefficients of the rational expectations solution are higher-order polynomials in the unrestricted Euler equation parameters and VAR coefficients. Hence, there are multiple solutions to the cross-equation restrictions, with the number of solutions rapidly increasing in the dimensions of the VAR. This multiplicity problem – thus far ignored by the literature – raises two important questions. First, do the solutions differ in their economic interpretation? Second, since each of the solutions implies different constraints on the ML estimation, which one should be selected?

The paper shows that the multiple rational expectations solutions can be classified according to two types that differ fundamentally in their economic interpretation. The first solution obtains by simply lagging the Euler equation in (1) by one period and rewriting it in VAR form. In this case, the cross-equation restrictions reduce to single-equation restrictions that are independent of the VAR specification and can be estimated via Ordinary Least Squares (OLS). However, OLS in this context is compatible with the regularity conditions of ML only under the strong assumption that the theory holds exactly in the data; i.e. $u_t = 0$ for all t . I refer to this solution as the *pure theory solution* because it amounts to testing a pure form of the theory that does not allow for any unexplained deviations (e.g. measurement errors).

All other solutions are of the second type and obtain under the weaker, presumably

more realistic assumption that the theory holds only up to time-varying yet unpredictable deviations; i.e. $u_t \neq 0$ with $E_{t-1}u_t = 0$. These solutions – henceforth called *approximate theory solutions* – are highly nonlinear functions of the Euler equation parameters and the remaining VAR coefficients. The non-linear character and the potentially large number of solutions means that computing the ML subject to each solution and then selecting the one with the highest likelihood quickly becomes impracticable. Instead, I show that the multiplicity problem arising in this case can be circumvented altogether by reformulating the cross-equation restrictions as constraints on the VAR coefficients of the forcing variable x_t . There exists only one mapping from these coefficients to the unrestricted parameters and thus, ML estimation boils down to a standard constrained optimization problem.

The proposed estimation strategy strongly contrasts with an alternative, widely employed method that consists of expressing the Euler equation and the VAR equations for the exogenous variables as a dynamic general equilibrium (DGE) system and restricting the ML estimation to yield a *unique stable* rational expectations solution.¹ In doing so, this alternative method implicitly avoids the multiplicity problem because it only considers approximate theory solutions and restricts the estimates such that all but one solution can be discarded on grounds of non-stationarity. The problem is, however, that the VAR component of the DGE system is only an approximation of market expectations rather than a structural description of the true dynamics. Hence, imposing uniqueness makes little economic sense and amounts to unnecessarily constraining the estimation, which can lead to severe misspecification in case the likelihood is maximized for a combination of parameters that implies multiple stable solutions.²

To illustrate the severity of imposing uniqueness, I apply the proposed estimation strategy to the hybrid New Keynesian Phillips Curve (NKPC) of Gali and Gertler (1999). The theory implies a log-linear Euler equation as in (1) that links current inflation to expected future inflation, past inflation and current real marginal cost. Using quarterly U.S. data for the sample period 1960-1997 and measuring real marginal cost with labor income share, my ML estimates indicate that the hybrid NKPC cannot be rejected by a conventional likelihood ratio test once we allow for uncorrelated deviations u_t . In addition, labor income share enters significantly into the model and forward-looking behavior is the predominant determinant of inflation. By contrast, if the same ML estimator is constrained additionally to yield a unique stable solution, the results change dramatically: the backward-looking

¹Examples that employ this alternative method are Fuhrer and Moore (1995), Jondeau and Le Bihan (2001) or Fuhrer and Rudebusch (2004).

²Some studies refer to this unique stable method as "*Full Information Maximum Likelihood*" (FIML) because it consists of specifying a dynamic equation for every variable of the system. In my view, this terminology is misleading, however, as it confounds the idea of estimating an Euler equation conditional on a VAR approximation of expectations with the idea of estimating a fully specified (i.e. full information) DGE model where each equation is derived from structural underpinnings.

inflation component of the hybrid NKPC becomes equally important than the forward-looking component; real marginal cost enters insignificantly; and the model as a whole is rejected. The example of the hybrid NKPC thus provides a telling illustration that imposing uniqueness can severely constrain the estimates and lead to wrong conclusions.

The results in this paper coincide by and large with the General Method of Moments (GMM) estimates of Gali and Gertler (1999).³ At the same time, they strongly contrast with ML estimates of Fuhrer and Moore (1995) and others who proxy real marginal cost with the output gap and overwhelmingly reject various forward-looking versions of the NKPC. This is interesting because Gali and Gertler’s findings have been challenged recently on account of the poor finite-sample performance of GMM. The close match with the supposedly more robust ML estimates here suggests, however, that the fit of the hybrid NKPC is not a question of estimation method but rather depends on how real marginal cost is measured. To further assess this issue, I reestimate the hybrid NKPC conditional on real marginal cost being proxied with the output gap. Analogous to Fuhrer and Moore, I find that in this case, the model is strongly rejected and the data attributes much less importance to the forward-looking component. Hence, the main empirical difficulty is not in explaining inflation with a forward-looking pricing model, but in reconciling the behavior of output with the behavior of real marginal cost. Moreover and contrary to the labor income share case, I find that the results conditional on the output gap are not affected by the uniqueness condition because the ML estimates by themselves already imply a unique stable rational expectations solution. This explains why my results are very similar to Fuhrer and Moore’s who impose uniqueness in their estimation and goes to show that the uniqueness condition unnecessarily constrains the estimates only if the likelihood surface implied by the data is maximized in a region with multiple stable solutions.

The paper is organized as follows. Section 2 considers a stylized example to discuss the derivation of cross-equation restrictions and the multiplicity problem. Section 3 illustrates the pitfalls of restricting the estimation to a unique stable solution. Section 4 applies the proposed estimation to the hybrid NKPC and places the results in the context of the existing literature. Section 5 concludes.

2 Cross-equation restrictions and the multiplicity problem

To illustrate the multiplicity problem, consider a purely forward-looking version of the Euler equation in (1)

$$(2) \quad y_t = aE_t y_{t+1} + bx_t + u_t.$$

³Sbordone (2002, 2004) also finds results similar to Gali and Gertler (1999) based on alternative minimum distance estimators.

Suppose furthermore that the dynamics of the forcing variable x_t are well-described by a bivariate VAR(1) of the form

$$(3) \quad \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} e_{x,t} \\ e_{y,t} \end{bmatrix}.$$

This process can be expressed more compactly as $\mathbf{z}_t = \mathbf{M}\mathbf{z}_{t-1} + \mathbf{e}_t$, where $[\mathbf{e}_t \ \mathbf{e}_{t+k}]' \sim (\mathbf{0}, \Sigma)$ with $\Sigma = \mathbf{0}$ for all $k \neq 0$, and where all roots of the matrix \mathbf{M} lie inside the unit circle. All of the results are derived with respect to this stylized example. However, the different propositions that follow from the derivations also hold for the more general Euler equation in (1) and for larger VAR processes, with the appendix providing the formal proofs.

2.1 Rational expectations cross-equation restrictions

Under rational expectations, the Euler equation in (2) implies cross-equation restrictions between the structural parameters a , b and the VAR coefficients. Following Sargent (1979), these restrictions can be derived in three steps.⁴ First, assuming that the information contained in \mathbf{z}_t is a subset of the market's full information set \mathbf{Z}_t and that the bivariate VAR in (3) provides a good approximation of the dynamics for \mathbf{z}_t , the law of iterated expectations implies

$$\begin{aligned} E_t y_{t+i} &= h_y \mathbf{M}^i \mathbf{z}_t \\ E_t x_{t+i} &= h_x \mathbf{M}^i \mathbf{z}_t, \end{aligned}$$

with $h_y = [0 \ 1]$ and $h_x = [1 \ 0]$.⁵ Second, these expectation formulas are used to rewrite the Euler equation in (2) conditional on information \mathbf{z}_{t-1} as

$$h_y \mathbf{M} \mathbf{z}_{t-1} = ah_y \mathbf{M}^2 \mathbf{z}_{t-1} + bh_x \mathbf{M} \mathbf{z}_{t-1} + E[u_t | \mathbf{z}_{t-1}].$$

Third, under the assumption that u_t is an i.i.d. error term, rational expectations implies $E[u_t | \mathbf{z}_{t-1}] = 0$.⁶ Since this relationship needs to hold for general \mathbf{z}_{t-1} , we obtain

$$(4) \quad h_y \mathbf{M} = ah_y \mathbf{M}^2 + bh_x \mathbf{M},$$

⁴Sargent derives these cross-equation restrictions in the context of a finite version of the expectations theory of interest rates. The steps used to derive the cross-equation restrictions for the Euler equations here are exactly the same.

⁵The law of iterated expectations says that for $\mathbf{z}_t \subseteq \mathbf{Z}_t$, we can write

$$E[E_t x_{t+i} | \mathbf{z}_t] = E[E x_{t+i} | \mathbf{Z}_t] | \mathbf{z}_t = E[x_{t+i} | \mathbf{z}_t].$$

Practically, this says that an econometrician's best estimate of market expectations about some variable x_{t+i} given \mathbf{Z}_t is equal to the econometrician's forecast from information \mathbf{z}_t .

⁶The assumption that u_t is uncorrelated is not innocuous. For example, if the econometrician assumed that the error term was serially correlated, then rational expectations would no longer imply $E[u_t | \mathbf{z}_{t-1}] = 0$. This, in turn, would affect the form of the cross-equation restrictions. I return to this issue at the end of the paper.

or written more explicitly

$$\begin{aligned} m_{yx}(1 - a(m_{xx} + m_{yy})) &= bm_{xx} \\ m_{yy}(1 - am_{yy}) - am_{xy}m_{yx} &= bm_{xy}. \end{aligned}$$

The 2 equations characterize the cross-equation restrictions between the structural parameters a , b of the theory and the VAR parameters m_{xx} , m_{xy} , m_{yx} , m_{yy} . This type of restrictions are what Hansen and Sargent (1980) call the hallmark of rational expectations models. Intuitively, they arise because the VAR forecasts of x_t and y_t must be consistent with the dynamic relationship between the two variables as predicted by the theory.

2.2 Maximum Likelihood estimation and multiplicity

Under the assumption that the VAR error terms \mathbf{e}_t are multivariate normal, the log-likelihood for sample $\{\mathbf{z}_t\}_{t=1}^T$ can be expressed as⁷

$$(5) \quad \mathcal{L} = -\frac{nT}{2} \log(2\pi) - \frac{T}{2} [\log[\det(\Sigma)] + n]$$

with $n = 2$. ML estimation of the Euler equation in (2) based on information \mathbf{z}_t therefore involves minimizing $\det(\hat{\Sigma})$, where

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t' \text{ with } \hat{\mathbf{e}}_t = \mathbf{z}_t - \hat{\mathbf{M}}\mathbf{z}_{t-1},$$

subject to the cross-equation restrictions in (4).

To implement this constrained optimization problem, Sargent (1979) uses the cross-equation restrictions to express the coefficients of the VAR equation for y_t (i.e. the rational expectations solution) as a function of the parameters of the Euler equation and the remaining (free) VAR coefficients. For the present case, this means transforming (4) into two explicit functions of the form

$$(6) \quad m_{y\cdot} = f(a, b, m_{x\cdot}),$$

where $m_{y\cdot} \equiv \{m_{yx}, m_{yy}\}$ and $m_{x\cdot} \equiv \{m_{xx}, m_{xy}\}$. The log-likelihood in (5) is then maximized with respect to a , b and $m_{x\cdot}$.

Sargent's approach would be straightforward if transforming (4) into an explicit function of the form given by (6) was easy. Unfortunately, this is far from trivial because even for low-dimensional VAR processes, there exist multiple solutions $m_{y\cdot}$ that satisfy the cross-equation restrictions. In fact, for the stylized example considered here, there are three different solutions. The first solution is

$$(7) \quad m_{yx} = -b/a, m_{yy} = 1/a.$$

⁷See Hamilton (1994) for a derivation.

The second and the third solution take the form

$$(8) \quad m_{yx} = \frac{m_{xx}}{m_{xy}} m_{yy}, m_{yy} = \frac{(1 - am_{xx}) - \sqrt{(1 - am_{xx})^2 - 4abm_{xy}}}{2a}$$

$$(9) \quad m_{yx} = \frac{m_{xx}}{m_{xy}} m_{yy}, m_{yy} = \frac{(1 - am_{xx}) + \sqrt{(1 - am_{xx})^2 - 4abm_{xy}}}{2a}.$$

This multiplicity problem of Sargent's approach is not an artifact of the simple VAR specification or the purely forward-looking Euler equation of the present stylized example. To the contrary, the number of mappings $m_{y \cdot} = f(a, b, m_{-y \cdot})$ that satisfy (4) increases with the dimension of the VAR process. For example, if I use a bivariate VAR with 3 lags instead of one (as in the application of Section 4), the total number of different solutions increases to 7. This raises two important questions. First, do the solutions differ in their economic interpretation? Second, since each of the solutions implies different constraints on the ML estimation (and thus a different likelihood), which one should be selected? The following provides answers to both questions.

2.3 Interpreting the different solutions

Consider solution (7) first. It could have been obtained by simply lagging the variables of the Euler equation by one period and rewriting it as

$$y_t = -\frac{b}{a}x_{t-1} + \frac{1}{a}y_{t-1} - \frac{1}{a}u_{t-1} + \varepsilon_t,$$

where $\varepsilon_t = y_t - E_{t-1}y_t$ is defined as a rational expectations forecast error. In this case, the rational expectations solution for y_t is simply a function of the structural parameters a , b but does not depend on the coefficients of other VAR equations. Hence, the cross-equation restrictions reduce to single-equation restrictions, which suggests equation-by-equation regression via Ordinary Least Squares (OLS). However, such an estimation approach is compatible with the present ML framework only if $u_{t-1} = 0$. To see this, remember that an important regularity condition behind OLS is $E(e_{yt}\mathbf{z}_{t-1}) = 0$. But this condition is fulfilled only if $u_{t-1} = 0$, otherwise we have $e_{yt} = -\frac{1}{a}u_{t-1} + \varepsilon_t$, which is correlated with information \mathbf{z}_{t-1} . The first rational expectations solution in (7) thus imposes the additional constraint that the theory does not contain any unexplained stochastic error term.⁸ This amounts to estimating and testing a very stringent form of the theory that is reminiscent of efficient markets regressions in the financial literature as exemplified by Roll (1969). In analogy to the pure expectations hypothesis of the term structure of interest rates, I refer to this solution as the *pure theory solution*.⁹

⁸Note that the pure theory solution could also have been obtained by postmultiplying each term of the cross-equation restrictions in (4) by M^{-1} . But as is obvious from the derivation of these restrictions worked out above, this operation is admissible for general \mathbf{z}_t if and only if $u_t = 0$.

⁹The pure expectations theory of interest rates says that term premia between short and long bond yields (i.e. deviations from the theory) are zero. See King and Kurmann (2002) for a recent discussion.

The economic interpretation and econometric implications of the second and third solution to the cross-equation restrictions in (8)-(9) are fundamentally different. By contrast to the pure theory solution, the coefficients m_y are now a function of the structural parameters a, b as well as the VAR coefficients $m_x \equiv \{m_{xx}, m_{xy}\}$. Hence, the characteristics of both of these solutions depend crucially on the econometrician's specification of market expectations (i.e. the VAR process). Furthermore, both solutions imply a *singular* VAR matrix \mathbf{M} . To see this, rewrite \mathbf{M} with the restrictions in (8)-(9) imposed

$$\mathbf{M} \equiv \begin{bmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{bmatrix} = \begin{bmatrix} m_{xx} & m_{xy} \\ \frac{m_{xx}}{m_{xy}} m_{yy} & m_{yy} \end{bmatrix}.$$

The eigenvalue polynomial of this matrix equals $\det(\mathbf{M} - \lambda \mathbf{I}) = \lambda[\lambda - (m_{yy} + m_{xx})] = 0$. Hence, the eigenvalues of M are $\lambda_1 = 0$ and $\lambda_2 = m_{yy} + m_{xx}$, or rewritten in terms of the second and third solution

$$(10) \quad \lambda_1 = 0, \lambda_2 = \frac{(1 + am_{xx}) + \sqrt{(1 - am_{xx})^2 - 4abm_{xy}}}{2a}$$

$$(11) \quad \lambda_1 = 0, \lambda_2 = \frac{(1 + am_{xx}) - \sqrt{(1 - am_{xx})^2 - 4abm_{xy}}}{2a}.$$

Since $\lambda_1 = 0$, \mathbf{M} is singular and the cross-equation restrictions in (4) cannot be postmultiplied by \mathbf{M}^{-1} as is possible for the pure theory solution above (see footnote 8). Hence, ML estimation subject to either the second or the third solution amounts to testing a weaker form of the theory that allows for an unpredictable stochastic error term u_t . Accordingly, I refer to this type of solution as the *approximate theory solution*.

All of these results also hold for the more general case and are summarized by the following proposition.

Proposition 1 *The rational expectations solutions to the cross-equation restrictions can be classified according to two types. The pure theory solution is independent of assumptions about market expectations and implies that no unexplained deviations from the theory exist; i.e. $u_t = 0$. The approximate theory solutions depend on assumptions about market expectations. They imply that the VAR companion matrix \mathbf{M} is singular and therefore allow for unexplained deviations from the theory; i.e. $u_t \neq 0$.*

Proof: see appendix.

2.4 Circumventing the multiplicity problem

The difference in economic interpretation between the pure theory solution and the approximate theory solution provides a first natural selection criteria. If the econometrician is interested in evaluating a strong form of the Euler equation that does not allow for unexplained deviations, then imposing the pure theory solution to the cross-equation restrictions

is appropriate and estimation is implemented via OLS. Under the presumably more realistic assumption that the theory holds up to an uncorrelated error term, ML estimation of the entire VAR subject to the approximate theory solutions is necessary. But this leaves open the question of how to select among the different approximate theory solutions (two in the present stylized example; many more if the VAR specification is more involved).

Theoretically, one could think of estimating the VAR subject to each of the different approximate theory solutions and then select the one with the highest ML value. However, such an approach is impracticable because the number of solutions increases rapidly in the dimensions of the VAR and because each of the solutions is a highly non-linear function of the unrestricted parameters. A much simpler strategy consists of expressing the VAR coefficients for the forcing variable x_t as a function of the structural parameters and the coefficients of the remaining VAR equations. For the stylized example of this section, the cross-equation restrictions in (4) are thus rewritten as

$$\begin{aligned} m_{xx} &= \frac{m_{yx}(1 - am_{yy})}{am_{yx} + b} \\ m_{xy} &= \frac{m_{yy}(1 - am_{yy})}{am_{yx} + b}. \end{aligned}$$

As is immediately apparent, this reformulation provides a unique mapping from $m_x \equiv \{m_{xx}, m_{xy}\}$ to the structural parameters a , b and the VAR coefficients of the endogenous variable $m_y \equiv \{m_{yx}, m_{yy}\}$; i.e. $m_x = f(a, b, m_y)$. Hence, rewriting the cross-equation restrictions as explicit solutions to the VAR coefficients of the forcing variable altogether circumvents the multiplicity problem of Sargent's original approach.

The following proposition summarizes this result for the more general case.

Proposition 2 *Cross-equation restrictions implied by Euler equations of the form $y_t = aE_t y_{t+1} + bx_t + cy_{t-1} + u_t$ have a single approximate theory solution in the VAR coefficients of the forcing variable x_t .*

Proof: see appendix.

3 Pitfalls of restricting the estimation to a unique stable solution

The proposed remedy to circumvent the multiplicity problem strongly contrasts with an alternative strategy, employed for example by Fuhrer and Moore (1995).¹⁰ This strategy consists of considering the Euler equation and the VAR equation(s) for the exogenous

¹⁰Other studies that employ this method are Fuhrer (1997), Jondeau and LeBihan (2001), or Fuhrer and Rudebusch (2004).

forcing variable as a structural DGE system and to restrict the system to yield a *unique stable* solution. The problem with this uniqueness condition is that it effectively imposes *additional* parameter constraints on the ML estimator, thus leading to potentially serious misspecification.

3.1 The stylized example reconsidered

To fully recognize the problem, reconsider the stylized example from above. The Euler equation in (2) and the VAR equation for the forcing variable x_t in (3) can be expressed as a DGE system of the form

$$(12) \quad \begin{aligned} y_t &= aE_t y_{t+1} + b x_t + u_t \\ x_t &= m_{xx} x_{t-1} + m_{xy} y_{t-1} + e_{xt}. \end{aligned}$$

This system can be cast in standard form (see Blanchard and Kahn, 1980)

$$(13) \quad \mathbf{A}E_t \mathbf{Y}_{t+1} = \mathbf{B} \mathbf{Y}_t + \mathbf{C} \mathbf{X}_t,$$

where $\mathbf{Y}_{t+1} = [x_{t+1} \ y_{t+1} \ x_t \ y_t]'$, $\mathbf{Y}_t = [x_t \ y_t \ x_{t-1} \ y_{t-1}]'$, $\mathbf{X}_t = [e_{xt} \ u_t]'$ and \mathbf{A} , \mathbf{B} and \mathbf{C} are matrices filled with the appropriate parameters. It is well known that (13) has a unique stable rational expectations solution (i.e. a unique stable approximate theory solution) if and only if the number of non-predetermined variables in \mathbf{Y}_t equals the number of generalized eigenvalues of (13) with modulus larger than one. This solution has the VAR representation $\mathbf{z}_t = \mathbf{M} \mathbf{z}_{t-1} + \mathbf{e}_t$, with \mathbf{M} satisfying the cross-equation restrictions $m_{y\cdot} = f(a, b, m_{-y})$.

For the example in (12), there are two non-predetermined variables (x_t and y_t) and the generalized eigenvalues are¹¹

$$\lambda_1 = 0, \lambda_{2,3} = \frac{(1 + am_{xx}) \pm \sqrt{(1 - am_{xx})^2 - 4abm_{xy}}}{2a}, \lambda_4 = \infty.$$

By definition, we have $|\lambda_1| < 1$ and $|\lambda_4| > 1$. The question of whether the system has a unique stable solution therefore boils down to whether the combination of parameters a , b , m_{xx} , m_{xy} is such that $|\lambda_2| > 1$ while at the same time $|\lambda_3| < 1$. If this is indeed the case, then forward iteration allows to eliminate the two unstable eigenvalues and what remains are the two stable eigenvalues, which are exactly the non-zero eigenvalues of \mathbf{M} of the two approximate theory solutions in (10). In other words, imposing uniqueness avoids the multiplicity problem discussed above by constraining a , b , m_{xx} , m_{xy} to a region where all but one of the approximate theory solutions can be discarded on grounds of non-stationarity.

¹¹The generalized eigenvalues are the solutions to $\det(\mathbf{B} - \lambda \mathbf{A}) = 0$. One of the eigenvalues is infinite because \mathbf{A} is singular. See King and Watson (1998) for details.

How do the inequality constraints imposed by this uniqueness condition matter for the ML estimation? As long as the combination of a , b , m_{xx} , m_{xy} that maximizes the likelihood implies a unique stable solution, the inequality constraints do not bind and thus, imposing uniqueness has no influence on the ML estimation. However, the ML estimates of a , b , m_{xx} , m_{xy} may as well be located in a region of the parameter space with more than one stable solution. In this case, the inequality constraints imposed by the uniqueness condition bind and prevent the ML estimation from reaching its true maximum.

To illustrate the effects of imposing uniqueness, consider a simple simulation experiment. Suppose that the Euler equation in (12) is indeed the true structural description of y_t and that the dynamics of x_t are well approximated by the bivariate VAR(1). Furthermore, let the different parameters take the values $a = 1$, $b = 0.2$, $m_{xx} = 0.6$ and $m_{xy} = 0.1$. For this combination, the DGE system formed by the two equations has two stable approximate theory solutions. The eigenvalues for the VAR matrix \mathbf{M} implied by these two solutions are $(0, 0.66)$ and $(0, 0.94)$, respectively. For the simulation, I generate a sample of 10000 observations based on the first of the two solutions (none of the conclusions below would change if I generated the sample from the second solution instead).

Now, consider an econometrician who is interested in estimating the parameter b and for some reason already knows that $a = 1$, $m_{xx} = 0.6$ and $m_{xy} = 0.1$. The solid lines of Figure 1 plot the log-likelihood implied by the two approximate theory solutions as a function of b . For $b < 0$, the non-zero eigenvalue of the second solution is larger than one in absolute value. This solution can be discarded on grounds of non-stationarity (i.e. it is not plotted) and we have a unique stable solution in the form of the first solution. For $0 < b < 0.4$, the non-zero eigenvalues of both solutions are smaller than one in absolute value and thus, we have two stable solutions.¹² However, only for the first solution is the likelihood maximized at the correct value of $b = 0.2$ (point A). The econometrician should therefore allow for multiple stable solutions and then select the first solution. This is exactly what the estimation approach developed in Section 2 does, with the only difference that it circumvents the problem of selecting among multiple stable solutions altogether by expressing the cross-equation restrictions in terms of the VAR coefficients of the forcing variable. By contrast, imposing uniqueness to avoid the multiplicity problem would severely constrain the estimation and result in a ML estimate of $\hat{b} = 0$ (point B), far away from the true ML value of $b = 0.2$.

Figure 1 also depicts the likelihood implied by the pure theory solution (dotted line). Note that for all values of b , this likelihood lies below the likelihood of the first approximate theory solution. This is due to the fact that the true economy is simulated under the assumption that there is an unpredictable error term u_t . Hence, the pure theory solution also results in misspecification with a ML estimate of $\hat{b} = 0$ (point C) and the distance between

¹²For $b \geq 0.4$, no real-valued approximate theory solution exists.

the likelihoods at $b = 0.2$ can be considered as a measure of the degree of misspecification.

3.2 Does it make sense to impose uniqueness?

One may of course ask whether constraining the ML estimation to a unique stable solution could be justified on economic grounds. While none of the studies that employ the uniqueness approach brings forward such a justification, I can think of one potential explanation as well as two counterarguments against it. In fact, one may be tempted to interpret the Euler equation and the VAR process for the forcing variable as a *structural* description of the economy and consider only *unique* rational expectations solutions as has become customary in most of the modern DGE literature.

The first counterargument against this justification is that the uniqueness condition by definition rules out sunspot equilibria, i.e. equilibria where belief shocks lead agents to revise their forecasts of endogenous variables.¹³ As Benhabib and Farmer (1999) and others argue, allowing for sunspot equilibria may be important to explain business cycle dynamics. Hence, it is not clear why one would want to exclude such equilibria, especially when the main interest is in estimating a single Euler equation.

The second, more fundamental counterargument concerns the assumption that a purely backward-looking VAR equation represents a structural description of the dynamics of the forcing variable. Modern macroeconomic theory of consumer and firm behavior emphasizes intertemporal decisions under uncertainty, which typically results in optimality conditions that contain expectational terms. The VAR process for the forcing variable should therefore be considered as a reduced-form approximation rather than a structural description of the true data-generating process. As such, imposing uniqueness makes no sense.

Using a VAR approximation instead of a fully specified structural model to describe the dynamics of the forcing variable has the advantage that the estimation of the Euler equation of interest is not conditioned on other fundamental assumptions about the economy. But this naturally raises the issue of how well the VAR approximates the true dynamics. The following proposition helps in assessing this question.

Proposition 3 *As long as the rational expectations solutions of the variables in the Euler equation have a state-space representation, their dynamics are described by an infinite-order VAR process that does not contain any other variables.*

Proof: see appendix.

The proposition justifies the use of reduced-form VAR equations for ML estimation of Euler equations. At the same, it raises the question of how the different cross-equation restrictions are affected by the truncation to obtain a finite-order VAR process and how,

¹³See Lubik and Schorfheide (2003) for a method to include these sunspot equilibria as rational expectations solutions to DGE models.

in turn, this truncation affects the ML estimates. While admittedly important, I leave this question open for further research.

4 Application to New Keynesian pricing

This section applies the estimation method developed in Section 2 to a popular New Keynesian pricing equation. The application has received great attention in the recent literature. Hence, the results reported here are interesting because they add to the current debate. More importantly, however, the application provides an example where imposing uniqueness greatly affects the ML estimates.

4.1 The hybrid New Keynesian Phillips curve

The NK pricing model that I consider has been proposed by Galí and Gertler (1999) as an extension of the sticky price model by Calvo (1983). Since the primary focus of this section is about estimation, I only discuss some of the key aspects of the model that will be important for the interpretation of the results.

Calvo’s sticky price model consists of a large number of imperfectly competitive but otherwise identical firms as proposed by Blanchard and Kiyotaki (1987). It is assumed that in every period, a random fraction $1 - \theta$ of these firms adjusts its price. The remaining fraction θ of firms must keep its price fixed, disregarding of the number of periods it has kept its price unchanged in the past. Hence, a firm keeps its price fixed for an average of $1/(1 - \theta)$ periods. An adjusting firm will set its new price rationally such as to maximize current and expected future profits (taking into account that it may not readjust for several periods). Galí and Gertler’s extension consists of imposing that among the price adjusting firms, only a fraction $1 - \omega$ is forward-looking and set its price optimally. The remaining fraction ω of adjusting firms is assumed to use a rule-of-thumb instead that consists of setting the new price equal to the average of last period’s new price, updated with last period’s inflation rate.

Aggregating over the different firms and log-linearizing, these assumptions result in an Euler equation that links current inflation π_t to expected future inflation $E_t\pi_{t+1}$, real marginal cost ψ_t and lagged inflation π_{t-1}

$$(14) \quad \pi_t = \gamma_f E_t \pi_{t+1} + \varphi \psi_t + \gamma_b \pi_{t-1} + u_t,$$

where u_t is an error term tagged on for empirical purposes.¹⁴ All variables denote percentage deviations from their respective steady states. The slope coefficients γ_f , γ_b and φ are

¹⁴Different interpretations of this error term u_t are possible. On the theoretical side, u_t may capture structural misspecifications of the model or approximation errors induced by the linearization (see Rotemberg and Woodford, 1997). On the empirical side, u_t may take into account errors that arise because of mismeasurement of either real marginal cost or inflation, or because the econometrician’s approximation of

functions of the underlying structural model parameters and are defined as

$$\begin{aligned}
 \varphi &= \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\phi} \\
 \gamma_f &= \frac{\beta\theta}{\phi} \\
 \gamma_b &= \frac{\omega}{\phi},
 \end{aligned}
 \tag{15}$$

with $\phi = \theta + \omega[1 - \theta(1 - \beta)]$, where β is the discount factor. As Sbordone (2002) points out, this definition of the slope coefficient φ is obtained under the strong assumption that firms can immediately and costlessly reallocate their capital stocks. Under her alternative assumption that relative capital stocks are fixed in the short-run, the form of the hybrid NKPC remains the same but the definition of φ changes to

$$\varphi = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\phi} \frac{(1-\alpha)}{(1-\alpha(1-\mu))},
 \tag{16}$$

where $0 < \alpha < 1$ is the capital share in the production function (assuming constant-returns-to-scale Cobb-Douglas technology), and $\mu > 1$ represents the elasticity of substitution among the differentiated goods.¹⁵ Since $(1-\alpha)/(1-\alpha(1-\mu)) < 1$, it is straightforward to show that under Sbordone's alternative assumption, a given combination of γ_f , γ_b , φ implies a smaller θ (and thus a smaller average degree of price fixity) as well as a smaller ω .

Gali and Gertler call (14) the 'hybrid' New Keynesian Phillips Curve (NKPC) because inflation is determined both by (forward-looking) expectations about future inflation as well as a (backward-looking) lagged inflation term.¹⁶ For the particular case where the fraction of rule-of-thumbers is zero, i.e. $\omega = 0$, the hybrid NKPC reduces to $\pi_t = \beta E_t \pi_{t+1} + \varphi \psi_t + u_t$, which is the log-linearized solution for inflation of the original Calvo model. This expression can be rewritten as

$$\pi_t = \varphi \sum_{i=0}^{\infty} \beta^i E_t \psi_{t+i} + u_t.$$

Hence, if all firms are rational profit maximizers, inflation is entirely determined by the present-value of future expected real marginal cost (plus the possible error term u_t).

market's expectations is inappropriate. As already noted previously, the i.i.d assumption of this error term is not innocuous. I return to this issue at the end of the paper.

¹⁵Similar redefinitions of φ are discussed in Gali, Gertler and Lopez-Salido (2003), Woodford (2003) as well as Eichenbaum and Fischer (2004).

¹⁶Note that for $\beta = 1$, we obtain the special case of $\gamma_f + \gamma_b = 1$ and Gali and Gertler's hybrid NKPC becomes very similar to earlier specifications. For example, Fuhrer and Moore (1995) set $\gamma_f = \gamma_b = 1/2$, which is derived from a relative wage contracting framework. Roberts (1997), in turn, proposes a hybrid pricing equation with $\gamma_f + \gamma_b = 1$ from a framework where lagged inflation appears because of adaptive expectations of part of the price setters. See Kozicki and Tinsley (2002) for further discussion.

4.2 The empirical controversy

Substantial controversy surrounds the empirical relevance of the hybrid NKPC, and in particular the importance of the forward-looking inflation term relative to the lagged inflation term.¹⁷ To a large part, this controversy revolves around the question of how to measure real marginal cost ψ_t . Initial studies by Fuhrer and Moore (1995) and Fuhrer (1997) among others proxy real marginal cost with the output gap and find that (i) the output gap is not a significant determinant of inflation; and (ii) the forward-looking component is empirically unimportant in explaining the strong persistence of inflation in the data. Hence, these studies conclude that NK pricing is largely at odds with observed inflation dynamics.

Gali and Gertler (1999) and Sbordone (2002) challenge this conclusion on grounds that the output gap is an inappropriate proxy and measure real marginal cost with labor income share (i.e. real unit labor cost) instead.¹⁸ The difference in results is striking: labor income share enters significantly into the NKPC, and forward-looking behavior becomes the predominant determinant of inflation dynamics. Gali and Gertler and Sbordone thus conclude that the main empirical difficulty is not in explaining inflation with a forward-looking pricing model, but in reconciling the behavior of output with the behavior of real marginal cost.

More recently, however, a number of papers have focused on the choice of estimation technique as an explanation for the difference in results in the aforementioned studies. In fact, both Fuhrer and Moore (1995) and Fuhrer (1997) estimate their Euler equation via ML. By contrast, Gali and Gertler use a GMM estimator.¹⁹ This GMM estimator does not directly impose cross-equation restrictions and has been criticized on account of its poor finite-sample properties.²⁰ Lindé (2002) as well as Fuhrer and Rudebusch (2004) illustrate

¹⁷The relative weight of the forward- vs. the backward-looking component has important consequences for optimal monetary policy. See for example Rotemberg and Woodford (1997); or Clarida, Gali and Gertler (1999).

¹⁸From an empirical point of view, using the output gap is problematic because the natural level of output, from which the output gap is constructed, is unobserved and ad-hoc proxies are likely to be ridden with considerable error. From a theoretical point of view, the output gap is equivalent to real marginal cost only under the strong assumption that capital stocks are fixed. By contrast, labor income share is directly observable and can be derived from optimal firm behavior under the presumably less restrictive assumption that production technology is log-linear in its inputs (see Bils, 1987).

¹⁹Sbordone (2002) employs a different estimation approach that consists of minimizing the distance between observed inflation and a theoretical inflation series implied by the model. This theoretical inflation series is conditional on an unrestricted VAR forecasting process and hence, Sbordone does not directly impose the cross-equation restrictions. To my knowledge, the finite-sample properties of this minimum distance estimator have not been explored, which is why I concentrate in the following on Gali and Gertler's GMM results.

²⁰The moment conditions of GMM are asymptotically equivalent to cross-equation restrictions under the assumption that the dynamics of the instruments used in GMM are appropriately described by a VAR (see appendix for details). These restrictions are, however, not directly imposed by the finite-sample moment

by means of Monte-Carlo simulations that in situations of weak identification GMM tends to overestimate the importance of the forward-looking component in hybrid Euler equations.²¹ In turn, Jondeau and Le Bihan (2003) argue – also based on Monte-Carlo simulations – that measurement errors and omitted variables can cause GMM to further overstate the forward-looking component. Concurrently, the same simulations indicate that the ML estimator appears to be much more robust. This difference in finite-sample properties matches well with results in Jondeau and Le Bihan (2001) who estimate the hybrid NKPC with both ML and GMM and measure real marginal cost first with the output gap and then with labor income share. They report a substantially larger degree of backwardness for the ML estimates than for the GMM estimates, independent of the measure for real marginal cost. One is thus tempted to conclude that Gali and Gertler’s forward-looking estimates of the hybrid NKPC have nothing to do with their use of labor income share as the measure for real marginal cost, but are due to finite-sample bias of their GMM estimates.

Such a conclusion is problematic for two reasons, however. First, the finite-sample properties of GMM and ML inferred from Monte-Carlo simulations very much depend on the set up of the experiment. There are no general results that say which estimation technique is best. Second, and more importantly for the purpose of the present paper, Jondeau and Le Bihan (2001) follow Fuhrer and Moore (1995) and impose the additional restriction on their ML estimates that the rational expectations solution for inflation is unique (as discussed in the previous section). By contrast and as the appendix shows, Gali and Gertler’s GMM estimator does not impose this uniqueness condition. It is thus natural to ask whether uniqueness restrictions rather than finite-sample bias are at the root of the conflicting results between GMM and ML. The following estimation assesses this possibility.

4.3 Maximum Likelihood estimator and data

To implement the ML estimation, we need to derive the cross-equation restrictions that the hybrid NKPC implies on a general VAR process in n variables and p lags thereof. Since the hybrid NKPC is of equivalent form than the general Euler equation in (1), these restrictions are analogous to the ones in the proof of Proposition 1 (see appendix) and take the form

$$(17) \quad h_\pi[\mathbf{M} - \gamma_f \mathbf{M}^2 - \gamma_b \mathbf{I}] = \varphi h_\psi \mathbf{M},$$

where \mathbf{M} is the $(np \times np)$ companion matrix of a VAR; \mathbf{I} is a $np \times np$ identity matrix; and h_π and h_ψ are $(np \times 1)$ selection vectors for π_t and ψ_t , respectively.

conditions.

²¹The problem of finite-sample bias of GMM estimates in the presence of weak identification (i.e. instruments that are insufficiently correlated with the instrumented variable in the model) is well documented in the literature. See for example Fuhrer, Moore and Schuh (1995), Stock, Wright and Yogo (2002) or Stock and Yogo (2003).

The pure theory solution to (17) obtains under the assumption that $u_t = 0$ for all t and takes the form

$$\pi_t = -\frac{\varphi}{\gamma_f}\psi_{t-1} + \frac{1}{\gamma_f}\pi_{t-1} - \frac{\gamma_b}{\gamma_f}\pi_{t-2} + \varepsilon_t.$$

In this case, ML estimation reduces to independent OLS estimations of the np different VAR equations, with the coefficients of the inflation equation restricted to one lag in real marginal cost and two lags in inflation.

The approximate theory solution to the cross-equation restrictions in (17) obtains under the assumption that $u_t \neq 0$ and represents a system of np equations. By Proposition 2, this system has a single solution in the np coefficients of the VAR equation for real marginal cost (i.e. the forcing variable of the hybrid NKPC) and can be expressed as

$$(18) \quad m_{\psi.} = f(\gamma_f, \gamma_b, \psi, m_{-\psi.}),$$

where $m_{\psi.}$ is the $np \times 1$ coefficient vector of the VAR equation for real marginal cost; and $m_{-\psi.}$ is the vector containing the $n(n-1)p$ coefficients of the remaining VAR equations. Analogous to Section 2, the log-likelihood of the VAR can be expressed as

$$\mathcal{L} = -\frac{nT}{2} \log(2\pi) - \frac{T}{2} [\log[\det(\Sigma)] + n].$$

Hence, ML estimation of the approximate theory solution boils down to minimizing $\det(\hat{\Sigma})$, subject to the restrictions in (18).

ML estimation of the unique stable approximate theory solution is implemented by maximizing the log-likelihood subject to the same cross-equation restrictions, but with the additional condition that the combination of estimated parameters lies in the region where all but one approximate theory solution can be discarded on grounds of non-stationarity. This condition is imposed by expressing the hybrid NKPC together with all the VAR equations except the one for inflation as a DGE system. As discussed in Section 3, this system can be cast in standard form as

$$(19) \quad \mathbf{A}E_t\mathbf{Y}_{t+1} = \mathbf{B}\mathbf{Y}_t + \mathbf{C}\mathbf{X}_t.$$

For there to be a unique stable approximate theory solution in inflation, the number of non-predetermined variables in \mathbf{Y}_t needs to equal the number of generalized eigenvalues with modulus larger than one. Hence, the uniqueness condition imposes inequality constraints on different combinations of parameters.²²

²²These uniqueness restrictions are imposed by estimating directly the eigenvalues of the DGE system instead of the parameters. In each estimation step, the unique stable solution is then computed numerically using the algorithm by King and Watson (1998).

Turning to the data, I consider both the case where real marginal cost is measured by labor income share and the case where it is measured by the output gap. Following Proposition 3, I assume that the dynamics of the respective real marginal cost measure and inflation are well approximated by a bivariate VAR process in the two variables. For the sake of comparison, all estimates are based on exactly the same U.S. data series that Gali and Gertler (1999) use in their study. The sample covers the period 1961:1-1997:4.²³ The rate of inflation π_t is represented by the first difference of the overall GDP deflator. Labor income share s_t is measured by nominal non-farm business unit labor cost deflated with the non-farm business GDP deflator. The output gap is constructed as $x_t \equiv y_t - y_t^*$, where y_t and y_t^* denote the observed level of output and the natural level of output, respectively. As in Fuhrer and Moore (1995), I measure output with non-farm business GDP and proxy the natural level y_t^* with a fitted linear trend.²⁴

4.4 Results

For both the VAR in labor income share and inflation and the VAR in the output gap and inflation, the Aikake Information Criterion selects an optimal lag number of three. Table 1 presents the unrestricted OLS coefficient estimates for the two processes. The large and highly significant coefficient estimates on s_{t-1} in the labor income share equation of Table 1a, on x_{t-1} in the output gap equation of Table 1b, and on π_{t-1} in both inflation equations illustrate the sluggish behavior of the three variables that we observe in the data. Most of the other coefficient estimates are insignificant, however. Especially, lags of inflation appear to have only weak predictive power for labor income share and the output gap.²⁵

Table 2 reports the ML estimates for the pure theory solution, the approximate theory solution and the unique stable solution of the hybrid NKPC. Standard errors (in parenthesis) are computed via the Berndt-Hall-Hall-Hansen (BHHH) method. All estimates are obtained using the simulated annealing algorithm by Goffe (1996). Under appropriate implementation, simulated annealing identifies the global maximum with probability one

²³I thank Jordi Gali for providing me with the data. Note that Gali and Gertler (1999) indicate 1960:1 as the starting date of their estimation. However, this date takes into account the 4 lags they use for their instruments. Hence, the true starting date of the estimation is 1961:1.

²⁴The citibase codes to construct the different variables are 'lbgdpu' for the overall GDP deflator, 'lbcpu' for nominal non-farm business unit labor cost and 'gbpuq' for real non-farm business GDP. All variables are transformed to logarithms and demeaned prior to estimation.

²⁵The insignificance of the lagged inflation terms is of some importance for the estimation because it provides some indication that γ_b may be identified. As Sargent (1987) shows, the forcing variable must be strictly exogenous for the coefficient on the backward-looking term of the hybrid Euler equation to be identified. In the present case, if π_{t-1} has predictive power for s_t and x_t , respectively, then it will be correlated with u_t and thus, γ_b cannot be estimated consistently (see Nason and Smith, 2003 for a more detailed discussion). I will not further investigate this issue, however, since the main objective of this empirical application is to assess the effects of imposing uniqueness on the ML estimation.

as the number of grid searches goes to infinity.²⁶ This property represents an important advantage over traditional numerical gradient algorithms because it is well-known that constrained log-likelihood surfaces often have multiple local maxima. Preliminary estimates indicate that multiple local maxima indeed exist in the present case and thus, gradient-based algorithms produce different results depending on the starting values. Finally, note that β is fixed to unity in all cases since it is imprecisely estimated, and that the other two structural parameters are constrained to lie within the range of theoretically admissible values; i.e. $0 \leq \theta \leq 1$ and $0 \leq \omega \leq 1$.²⁷

4.4.1 Pure theory solution

Part (a) of Table 2 reports the ML estimates for the pure theory solution of the hybrid NKPC. When real marginal cost is measured by labor income share, the estimated fraction of firms that keep their prices fixed in any given period equals $\hat{\theta} = 0.92$, while the estimated fraction of backward-looking price setters ω is virtually zero. Together with the value of $\beta = 1$, these estimates imply a NKPC that is purely forward-looking. However, the likelihood ratio test versus the unrestricted VAR(3) strongly rejects the model and labor income share enters insignificantly into the NKPC.

An analogous picture emerges when real marginal cost is measured by the output gap: the point estimates of θ and ω are at their higher and lower boundary, respectively (note, however, that θ is very imprecisely estimated). Furthermore, the likelihood ratio test rejects the model even more strongly than in the labor income share case. In sum, the pure theory solution is decisively rejected in the data, disregarding of whether real marginal cost is measured by labor income share or the output gap. At the same time, the assumption that the hybrid NKPC holds exactly in the data is arguably too strong for it to be taken seriously.

4.4.2 Approximate theory solution

Part (b) of Table 2 reports the ML estimates of the approximate theory solution; i.e. the rational expectations of inflation that allows for uncorrelated deviations from the hybrid NKPC. As is immediately apparent, the results strongly contrast with the ones obtained for the pure theory solution. Furthermore and importantly in light of the empirical controversy discussed above, the estimates vary greatly depending on whether real marginal cost is measured by labor income share or the output gap.

²⁶A MATLAB version of the simulated annealing algorithm is available from the author upon request. The code is an adapted version of Goffe's (1996) FORTRAN code. See Goffe (1996) and Judd (1999) for further discussion on simulated annealing.

²⁷Setting β slightly smaller than 1 does not change the results. Furthermore, if β is estimated within the admissible bounds of $0 \leq \beta \leq 1$, the ML estimator pushes β towards 1.

Consider the labor income share case first. The likelihood ratio of 4.844 implies a p-value of 56% and thus, the hybrid NKPC cannot be rejected at conventional significance levels once one allows for uncorrelated deviations. Furthermore, for the standard specification of the slope coefficient φ in (15), the estimated frequency of price adjustment is $\hat{\theta} = 0.848$, implying an average price fixity of $1/(1 - \hat{\theta}) = 6.67$ quarters. This is too high compared with evidence from micro studies where firms are reported to change their price on average every 2.5 to 4 quarters.²⁸ However, using Sbordone's (2002) alternative definition of φ in (16) with $\alpha = 1/3$ and $\mu = 10$ (a 11% markup), the estimate drops to $\hat{\theta} = 0.671$, thus implying a more reasonable average price fixity of roughly 3 quarters. In turn, the fraction of backward-looking rule-of-thumbers is estimated at $\hat{\omega} = 0.050$ (respectively $\hat{\omega} = 0.040$ for the alternative definition of φ). Most firms therefore seem to behave rationally and set new prices so as to maximize the discounted present value of current and expected future profits. Together with $\beta = 1$, these estimates imply a hybrid NKPC of the form

$$\pi_t = \frac{0.944}{(0.147)} E_t \pi_{t+1} + \frac{0.024}{(0.015)} s_t + \frac{0.056}{(0.146)} \pi_{t-1} + \hat{u}_t.$$

This is surprisingly close to the GMM based results of Gali and Gertler (1999), who report a probability of price fixity θ between 0.803 and 0.838, and a fraction of backward-looking price-setters ω between 0.244 and 0.522 (depending on the moment specification). These estimates imply a lagged inflation coefficient γ_b that ranges from 0.233 to 0.383, which is somewhat larger than in the present case. However, the relatively large standard error of 0.146 for the present estimate of γ_b indicates that the degree of backwardness is imprecisely estimated and we cannot reject that the NKPC is purely forward-looking nor that γ_b falls in the range of Gali and Gertler's GMM estimate. Likewise, the ML estimate of the slope coefficient on labor income share of $\hat{\varphi} = 0.024$ is close to Gali and Gertler's (between 0.009 and 0.027) and significantly different from zero at the 89% level.²⁹ The present ML estimates therefore corroborate Gali and Gertler's main conclusion: once real marginal cost is measured by labor income share and once one allows for uncorrelated deviations, forward-looking behavior becomes predominant and the NKPC provides a good approximation of U.S. inflation dynamics.³⁰

The picture is much different for the case where real marginal cost is measured by the

²⁸See Taylor (1998) and Wolman (2000) or Bils and Kleenow (2004).

²⁹Interestingly, the present results are also very much similar to ML estimates of the hybrid NKPC obtained within a fully specified, structural DGE model. See for example Ireland (2001) or Smets and Wouters (2003).

³⁰The ML estimates here simply confirm Gali and Gertler's main conclusion, which does not mean that their GMM estimates do not suffer from finite-sample problems. In fact, Ma (2002) finds that bias-robust continuous-updating estimates of Gali and Gertler's hybrid NKPC attribute a much more important role to the backward-looking inflation component. His parameters are very imprecisely estimated, however, which makes inference about the degree of backwardness very difficult and indicates that weak identification is indeed a problem for method of moments estimators.

output gap. First, the likelihood ratio test rejects the hybrid NKPC at a high significance level. Second, the degree of backwardness is much more important than in the labor income share case, resulting in a hybrid NKPC of the form

$$\pi_t = \underset{(0.955)}{0.552} E_t \pi_{t+1} + \underset{(4.022)}{0.123} x_t + \underset{(0.587)}{0.448} \pi_{t-1} + \hat{u}_t.$$

The slope coefficients indicate that the forward- and the backward-looking part of inflation are about equally important. Furthermore, the coefficient on the output gap is not significantly different from zero.³¹ These results are very much similar to the ones reported by Fuhrer and Moore (1995) and Fuhrer (1997). They underline that the fit of the hybrid NKPC and more particularly the importance of the forward-looking component depends crucially on how real marginal cost is measured.

4.4.3 Unique stable solution

Part (c) of Table 2 reports the ML estimates when the parameters are constrained to yield a unique stable solution for inflation. For the case where real marginal cost is measured by labor income share, the estimates change dramatically. In particular, the fraction of backward-looking price setters jumps to $\hat{\omega} = 0.740$, which results in a hybrid NKPC with roughly equal coefficients on the forward- and the backward-looking part, and a labor income share estimate that is insignificantly different from zero³²

$$\pi_t = \underset{(2.122)}{0.542} E_t \pi_{t+1} + \underset{(0.092)}{0.003} s_t + \underset{(1.541)}{0.458} \pi_{t-1} + \hat{u}_t.$$

The likelihood ratio implied by these estimates is much higher and thus, the model as a whole is strongly rejected. These results are very similar to the ones reported by Jondeau and Le Bihan (2001) and suggest that indeed, they obtain a much more important degree of backwardness in the labor income share case because they constrain their estimates to yield a unique stable solution.³³ Hence, uniqueness restrictions rather than finite-sample differences seem to explain the conflicting results obtained from GMM and ML.

Interestingly, for the case where real marginal cost is measured with the output gap, none of the estimates are affected by the uniqueness condition. To understand this result,

³¹The large BHHH standard errors indicate that all of the slope coefficients are very imprecisely estimated. If these standard errors are instead inferred via the delta method from the standard errors of the structural coefficients, the degree of imprecision decreases greatly. While not central to this paper, this difference in standard errors suggests that numerical approximations of the information matrix can result in substantial inaccuracy.

³²Note that the BHHH standard errors on γ_f and γ_b are again very large. If they were inferred using the delta method, they would be much smaller (see footnote 31 above).

³³Jondeau and Le Bihan (2001) estimate the hybrid NKPC conditional on a VAR in three rather than two variables (their third variable being the short nominal rate) and their sample period is shorter. Nevertheless, I strongly suspect from the results reported here that their estimated degree of backwardness is due to their uniqueness constraint.

express the hybrid NKPC and the VAR(3) for the output gap as a DGE system as in (19). This system has two non-predetermined variables and, when evaluated at the ML estimates for the output gap in part (b) of Table 2, exactly two generalized eigenvalues with modulus larger than one (not reported here). Hence, the ML estimates for the output gap case already lie in the region of the parameter space with only one stable solution. This explains why the ML estimates for the output gap case resemble so much the ones reported by Fuhrer and Moore (1995) and Fuhrer (1997) even though they impose uniqueness in their estimation. By contrast, if the same DGE system is evaluated at the ML estimates for the labor income share case, there is only one generalized eigenvalue with modulus larger than one. In other words, there are multiple stable approximate theory solutions for inflation at the maximum, which is why imposing uniqueness has such a substantial impact in the labor income share case.

In sum, the estimates in this last subsection of the paper illustrate that imposing uniqueness can severely constrain the ML estimation of Euler equations. Whether these constraints are important or not depends on the likelihood surface implied by the data. If the likelihood is maximized for a combination of parameters that imply by themselves a unique stable solution, imposing uniqueness has no effect. However, if the likelihood is maximized at a point with multiple stable solutions, imposing uniqueness affects the estimation.

5 Conclusion

The paper improves on the existing literature along three important dimensions. First, it uncovers that there are multiple rational expectations solutions to the cross-equation restrictions of Sargent's (1979) original ML estimator and that these solutions can be classified according to whether they allow for uncorrelated deviations from the theory or not. Second, for the presumably more realistic case that admits uncorrelated deviations (i.e. the approximate theory solution), the paper proposes a novel estimation strategy that circumvents the multiplicity problem by solving the cross-equation restrictions in terms of the coefficients of the forcing variable of the theory. Third, the paper contrasts the proposed strategy to an alternative, widely employed method that consists of restricting the ML estimates to yield a unique stable rational expectations solution. I argue that imposing such a uniqueness condition makes little economic sense and illustrate by means of an application to Gali and Gertler's hybrid NKPC that the restrictions thus imposed can severely constrain the ML estimates.

In my view, an important avenue to investigate concerns the cross-equation restrictions that arise for more complicated models. As mentioned in the paper, the proposed estimation strategy applies to cross-equation restrictions derived under the assumption that deviations from the theory u_t are uncorrelated with past information. But such an assumption could be

overly restrictive, for example in the presence of systematic measurement errors.³⁴ Likewise, we may want to estimate Euler equations that involve multiple leads and lags of the different variables. The cross-equation restrictions that result from such specifications are more complicated because they have to be derived conditional on information earlier than $t - 1$.³⁵ Hence, there may no longer be a single solution in the coefficients of the forcing variable. Whether and how we can circumvent the multiplicity problem in these more general settings is the topic of future research.

³⁴For example, Roberts (2001) reports that once u_t is specified as a first-order autoregressive process, his GMM estimate of the marginal cost coefficient φ becomes negative, thus contradicting the theory.

³⁵Another reason why one might to derive the cross-equation restrictions conditional on information earlier than $t - 1$ is the identification problem of the backward-looking component mentioned in footnote 25.

A Appendix

A.1 Cross-equation restrictions and singularity of M

Consider the general Euler equation (1)

$$y_t = aE_t y_{t+1} + bx_t + cy_{t-1} + u_t,$$

and suppose that the dynamics of the forcing variable x_t are described by a stationary VAR process in n variables and p lags thereof

$$z_t = M_1 z_{t-1} + M_2 z_{t-2} + \dots + M_p z_{t-p} + e_{z,t},$$

where z_t is the n -variable vector of date t information and contains as a minimum the two variables of the Euler equation, x_t and y_t . For the sake of concreteness, let x_t and y_t take the first and second position in z_t , respectively. Analogous to the illustrative example in Section 2, this VAR process can be rewritten in companion form as

$$\mathbf{z}_t = \mathbf{M}\mathbf{z}_{t-1} + \mathbf{e}_t,$$

where $\mathbf{z}_t = [z_t \ z_{t-1} \dots z_{t-p+1}]'$ is the $np \times 1$ vector of relevant information; $\mathbf{e}_t = [e_{z,t} \ 0 \dots 0]'$ is the $(np \times 1)$ vector of rational expectations errors with $E[e_t | \mathbf{z}_{t-1}] = 0$; and \mathbf{M} is the $(np \times np)$ companion matrix given by

$$\mathbf{M} = \begin{bmatrix} M_1 & M_2 & \cdot & \cdot & \cdot & M_{p-1} & M_p \\ I_n & 0_n & \cdot & \cdot & \cdot & 0_n & 0_n \\ 0_n & I_n & \cdot & \cdot & \cdot & 0_n & 0_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0_n & 0_n & \cdot & \cdot & \cdot & I_n & 0_n \end{bmatrix}.$$

The $(n \times n)$ blocks M_i , $i = 1 \dots p$, contain the projection coefficients of z_{t-i} on z_t while I_n and 0_n are $(n \times n)$ identity and null matrices, respectively. Hence, M contains $n^2 p$ non-trivial coefficients that we can stack in a column vector $m = [m_1 \ m_2 \dots m_n] = \text{vec}([M_1 \ M_2 \dots M_p]')$, where m_j , $j = 1, 2 \dots n$ holds the np coefficients of the VAR equation for the j -th variable in z_t .

Under the assumption that the econometrician's information set is a subset of the agents' full information set (i.e. $\mathbf{z}_t \subseteq \mathbf{Z}_t$) and under the assumption that the error term u_t is unforecastable given information $t-1$ and before (i.e. $E[u_t | \mathbf{Z}_{t-1}] = 0$), the rational expectations cross-equation can be derived analogous to Section 2 as

$$(20) \quad h_y[\mathbf{M} - a\mathbf{M}^2 - c\mathbf{I}_{np}] = bh_x\mathbf{M},$$

where h_x and h_y are $(1 \times np)$ selection vectors. Given the supposed placement of x_t and y_t as the first and second element of \mathbf{z}_t , these selection vectors are defined as $h_x = [1 \ 0 \ 0 \dots 0]$ and $h_y = [0 \ 1 \ 0 \ 0 \dots 0]$.

To prove the singularity of \mathbf{M} under the approximate theory solution, note that a scalar λ and a vector t are said to be a (left) eigenvalue / eigenvector pair of \mathbf{M} if

$$t\mathbf{M} = \lambda t,$$

which can be rewritten as $(\mathbf{I}_{np} - \lambda\mathbf{M})t = 0$. Thus, as long as a nonzero vector t exists, an eigenvalue λ of \mathbf{M} is a number that satisfies the characteristic equation

$$\det(\mathbf{I}_{np} - \lambda\mathbf{M}) = 0.$$

Equivalently, this condition can be expressed as a np -th order polynomial in λ ; i.e. $\det(\mathbf{I}_{np} - \lambda\mathbf{M}) = \lambda^{np} + c_1\lambda^{np-1} + c_2\lambda^{np-2} + \dots c_{np-1}\lambda + c_{np} = 0$. For \mathbf{M} to be singular, there needs to be at least one zero eigenvalue.³⁶ Hence, one needs to show that $c_{np} = 0$ under the approximate theory solution. To do so, I first show that the particular structure of the companion matrix implies $c_{np} = \det(M_p)$. Second, I derive that $\det(M_p) = 0$ under the approximate theory solution.

For the first part, rewrite the argument of the characteristic equation as

$$[\mathbf{M} - \lambda\mathbf{I}_{np}] \equiv C = \begin{bmatrix} M_1 - \lambda I_n & M_2 & \cdot & \cdot & M_{p-2} & M_{p-1} & M_p \\ I_n & -\lambda I_n & \cdot & \cdot & \cdot & 0_n & 0_n \\ 0_n & I_n & \cdot & \cdot & \cdot & 0_n & 0_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & -\lambda I_n & 0_n & 0_n \\ \cdot & \cdot & \cdot & \cdot & I_n & -\lambda I_n & 0_n \\ 0_n & 0_n & \cdot & \cdot & 0_n & I_n & -\lambda I_n \end{bmatrix}.$$

Now, note that determinant for any $(r \times r)$ matrix A can be expressed as

$$\det(A) = \sum_{j=1}^r (-1)^{j+1} a_{r,j} \det(A_{r,j}),$$

where $a_{r,j}$ denotes the r -th row / j -th column element of A , and where $A_{r,j}$ denotes the $(r-1 \times r-1)$ matrix formed by deleting row r and column j from A . Applying this formula to the present case, I obtain

$$\det([\mathbf{M} - \lambda\mathbf{I}_{np}]) \equiv \det(C) = -\det(C_{np,n(p-1)}) + \lambda \det(C_{np,np}),$$

³⁶This follows directly from the Jordan decomposition $V\mathbf{M}V^{-1} = J$ since a zero eigenvalue implies $\det(J) = 0$ and $\det(J) = \det(\mathbf{M})$ by definition.

as all the other elements in the last row of $[\mathbf{M} - \lambda \mathbf{I}_{np}] \equiv C$ are zero. To derive the part c_{np} that does not contain any λ , I only need to further consider the first term of $\det(C)$, which I redefine as $-\det(C_{np,n(p-1)}) \equiv -\det(C^1)$ for notational convenience. Applying the same determinant formula again, I can write $-\det(C^1)$ as

$$-\det(C^1) = \det\left(C_{n(p-1),n(p-2)}^1\right) - \lambda \det\left(C_{n(p-1),n(p-1)}^1\right).$$

Again, one only needs to further consider the first term $\det\left(C_{n(p-1),n(p-2)}^1\right)$, which is redefined in similar fashion as $\det(C^2)$. Its determinant is

$$\det(C^2) = -\det\left(C_{n(p-2),n(p-3)}^2\right) + \lambda \det\left(C_{n(p-2),n(p-2)}^2\right).$$

This reduction can be done for total of $n(p-2) + 1$ times until the only element of $\det([\mathbf{M} - \lambda I]) = \det(C)$ left to consider is

$$-\det(C^{n(p-2)+1}) = -\det\left(\begin{bmatrix} M_1 - \lambda I_n & M_p \\ I_n & 0_n \end{bmatrix}\right).$$

Finally, it can be shown that for any matrix with the $(n \times n)$ blocks D, E, F ³⁷

$$\det\left(\begin{bmatrix} D & E \\ F & 0_n \end{bmatrix}\right) = -\det(EF').$$

Applying this result to the present case, one obtains

$$-\det(C^{n(p-2)+1}) = \det(M_p),$$

and thus, it must be the case that $c_{np} = \det(M_p)$.

Turning to the second part of the proof, write out the rows of the cross-equation restrictions in (20) as

$$\begin{aligned} & [h_{y,n} \ 0_n \dots 0_n] \left(\begin{bmatrix} M_1 & M_2 & \dots & M_p \\ I_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & I_n & 0_n \end{bmatrix} - a \begin{bmatrix} M_1 & M_2 & \dots & M_p \\ I_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & I_n & 0_n \end{bmatrix}^2 - c \begin{bmatrix} I_n & 0_n & \dots & 0_n \\ 0_n & I_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & 0_n & I_n \end{bmatrix} \right) \\ & = b[h_{x,n} \ 0_n \dots 0_n] \begin{bmatrix} M_1 & M_2 & \dots & M_p \\ I_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & I_n & 0_n \end{bmatrix}, \end{aligned}$$

³⁷This result is taken from *The Matrix Reference Manual* by Mike Brookes (www.ee.ic.ac.uk/hp/staff/dmb/matrix/).

where $h_{y,n}$ and $h_{x,n}$ are $(1 \times n)$ selection vectors. Rewriting h_y and h_x in this form highlights that only the first n rows of the cross-equation restrictions matter and thus, (20) can be reduced without loss of generality to

$$\begin{aligned} & h_{y,n} \begin{bmatrix} M_1 - a(M_1^2 + M_2) - cI_n & M_2 - a(M_1M_2 + M_3) & \dots & M_p - aM_1M_p \end{bmatrix} \\ = & bh_{x,n} \begin{bmatrix} M_1 & M_2 & \dots & M_p \end{bmatrix}. \end{aligned}$$

The last n equations of this expression are

$$h_{y,n} [I_n - aM_1] M_p = bh_{x,n} M_p,$$

or equivalently

$$(h_{y,n} [I_n - aM_1] - ch_{x,n}) M_p = 0.$$

There are two possible solutions to this condition. Either $v = (h_{y,n} [I_n - aM_1] - ch_{x,n}) = 0$ (i.e. v is a zero-vector), in which case we cannot say anything about the characteristics of M_p . Or, v is not a zero-vector, in which case v represents an eigenvector associated with a zero eigenvalue of M_p ; i.e. $\det(M_p) = 0$. Hence, the cross-equation restrictions imply a singularity on the VAR companion matrix M as long as $v \neq 0$.

To only part left in the proof is to examine the conditions for which $v = 0$; i.e. the conditions for which the cross-equation restrictions do not imply a singularity. Write out the rows of $v = (h_{y,n} [I_n - aM_1] - ch_{x,n})$ as

$$\begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_j \\ \dots \\ v_n \end{bmatrix} = \begin{bmatrix} -am_{xy,1} - c \\ 1 - am_{yy,1} \\ \dots \\ -am_{yj,1} \\ \dots \\ -am_{yn,1} \end{bmatrix},$$

which makes clear that $v = 0$ if and only if $m_{yx,1} = -c/a$, $m_{xx,1} = 1/a$ and $m_{yj,1} = 0$ for all other j . Furthermore, the condition $v = 0$ can be used to derive restrictions for the other coefficients m_y of the VAR equation for the endogenous variable. Consider the first n terms of the cross-equation restrictions above

$$h_{y,n} (M_1 - a(M_1^2 + M_2) - cI_n) = bh_{x,n} M_1,$$

which can be rearranged as

$$(h_{y,n} [I_n - aM_1] - bh_{x,n}) M_1 = h_{y,n} (aM_2 + cI_n),$$

or equivalently

$$vM_1 = h_{y,n} (aM_2 + cI_n).$$

The condition $v = 0$ therefore implies $h_{y,n} (M_2 + bI_n) = 0$, or written more explicitly $m_{yx,2} = 0$, $m_{yy,2} = -b/a$ and $m_{yj,2} = 0$ for all other j . Analogously, all the other $k = 2, 3 \dots p - 1$ sets of cross-equation restrictions can be expressed as

$$(h_{y,n} [I_n - aM_1] - bh_{x,n}) M_k = -h_{y,n} M_{k+1},$$

or equivalently

$$vM_k = -h_{y,n} M_{k+1}.$$

For $v = 0$, one obtains $m_{yj,k} = 0$ for all $j = 1, 2, \dots n$.

To summarize, the cross-equation restrictions result in a singular companion matrix \mathbf{M} for all cases but $v = 0$. This latter case implies coefficient restrictions of the form

$$m_{yx,1} = -b/a, m_{yy,1} = 1/a, m_{yx,2} = 0, m_{yy,2} = -c/a \text{ and } m_{yj,k} = 0 \text{ for all other } j \text{ and } k.$$

But these are exactly the restrictions of the *pure theory solution*. Hence, it is exactly the pure theory solution that does not imply a singular \mathbf{M} , which means that one can postmultiply each term of (20) by \mathbf{M}^{-1} such as to recover restrictions that hold under the condition $u_t = 0$ and that do not depend on the specification of the VAR process. This proves the first proposition.

A.2 Circumventing the multiple solutions problem

Consider the cross-equation restrictions (20)

$$h_y[\mathbf{M} - a\mathbf{M}^2 - c\mathbf{I}_{np}] = bh_x\mathbf{M},$$

and use the same definition of the companion matrix \mathbf{M} as before. Writing out these np restrictions equation by equation, we obtain

$$\begin{aligned}
m_{yx,1} - a(m_{yx,1}m_{xx,1} + m_{yy,1}m_{yx,1} + \sum_{j=3}^n m_{yj,1}m_{jx,1} + m_{yx,2}) &= bm_{xx,1} \\
m_{yy,1} - a(m_{yx,1}m_{xy,1} + m_{yy,1}m_{yy,1} + \sum_{s=3}^n m_{ys,1}m_{sy,1} + m_{yy,2}) - c &= bm_{xy,1} \\
m_{y3,1} - a(m_{yx,1}m_{x3,1} + m_{yy,1}m_{y3,1} + \sum_{s=3}^n m_{ys,1}m_{s3,1} + m_{y3,2}) &= bm_{x3,1} \\
&\dots \\
m_{yj,k} - a(m_{yx,1}m_{xj,k} + m_{yy,1}m_{yj,k} + \sum_{s=3}^n m_{ys,k}m_{sj,k}) &= bm_{xj,k} \\
&\dots \\
m_{yn,p} - a(m_{yx,1}m_{xn,p} + m_{yy,1}m_{yn,p} + \sum_{s=3}^n m_{ys,p}m_{sn,p}) &= bm_{xn,p}
\end{aligned}$$

where $j = 3, 4, \dots, n$ denotes the j -th variable in the vector of date t information z_t ; and $k = 2, 3, \dots, p$ denotes the k -th lag. According to Sargent's (1979) approach, the cross-equation restrictions are imposed by solving this system for the VAR coefficients of the equation for the endogenous variable y ; i.e.

$$m_{y\cdot} = f(a, b, c, m_{-y}).$$

Since several terms of $m_{y\cdot}$ appear in each equation non-linearly, these solutions turn out to be complicated higher-order polynomials. Therefore, multiple $m_{y\cdot}$ satisfy the approximate theory solution to the cross-equation restrictions. By contrast, it is straightforward to see that each equation only contains one VAR coefficient of the equation for the exogenous variable; i.e. $m_{xx,1}$ is the only one of these coefficients that enters the first equation; $m_{xy,1}$ is the only coefficient that enters the second equation; and so forth. Hence, there is a single approximate theory solution of the form

$$m_{x\cdot} = f(a, b, c, m_{-x}).$$

This proves the second proposition.

A.3 VAR approximation

Suppose that the dynamics of the variables of Euler equation, x_t and y_t , can be represented in state-space form

$$\begin{aligned}
z_t &= \begin{bmatrix} x_t \\ y_t \end{bmatrix} = H'S_t + V_t \\
S_t &= FS_{t-1} + U_t,
\end{aligned}$$

where S_t is a vector of state variables; V_t and U_t are vectors of uncorrelated error terms; and H' and F are corresponding loading matrices. According to the Kalman Filter, we know that the best linear forecast of the state vector equals³⁸

$$E_t S_{t+1} = F E_{t-1} S_t + K_t (z_t - H' E_t S_{t-1}),$$

where K_t is the Kalman gain matrix. In steady state, this gain is constant and we can rewrite the forecast as

$$E_t S_{t+1} = [I - (F - K H') L]^{-1} K z_t,$$

where L is the lag operator. Since $E_t z_{t+1} = H' E_t S_{t+1}$ by definition of the state-space representation, the best linear forecast of z_t is

$$E_t z_{t+1} = H' [I - (F - K H') L]^{-1} K z_t.$$

In other words, z_t has an infinite-order VAR representation in its own variables as long as it can be represented in state-space form. This proves the third proposition.

A.4 Deriving cross-equation restrictions from GMM moment conditions

Consider the hybrid NK pricing equation

$$\pi_t = \gamma_f E_t \pi_{t+1} + \varphi \psi_t + \gamma_b \pi_{t-1} + u_t,$$

and rewrite it as

$$\pi_t - \gamma_f \pi_{t+1} - \varphi \psi_t - \gamma_b \pi_{t-1} = u_t + \xi_{t+1},$$

where $\xi_{t+1} = \pi_{t+1} - E_t \pi_{t+1}$ is the expectations error about future inflation. Under rational expectations, this expectation error and the error term u_t are uncorrelated with information \mathbf{z}_{t-1} , i.e. $E[u_t + \xi_{t+1} | \mathbf{z}_{t-1}] = 0$. Hence, we obtain the following moment conditions

$$E[(\pi_t - \gamma_f \pi_{t+1} - \varphi \psi_t - \gamma_b \pi_{t-1}) \mathbf{z}'_{t-1}] = 0.$$

Now, assume that the evolution of the instruments \mathbf{z}_{t-1} is described by the VAR process

$$\mathbf{z}_t = \mathbf{M} \mathbf{z}_{t-1} + \mathbf{e}_t.$$

Hence, the moment conditions can be rewritten as

$$E[h_\pi \mathbf{M} \mathbf{z}_{t-1} \mathbf{z}'_{t-1} - \gamma_f h_\pi \mathbf{M}^2 \mathbf{z}_{t-1} \mathbf{z}'_{t-1} - \varphi h_\psi \mathbf{M} \mathbf{z}_{t-1} \mathbf{z}'_{t-1} - \gamma_b h_\pi \mathbf{z}_{t-1} \mathbf{z}'_{t-1}] = 0,$$

³⁸See Hamilton (1994, page 390 ff) for details.

or equivalently

$$h_\pi \mathbf{M} \Sigma - \gamma_f h_\pi \mathbf{M}^2 \Sigma - \varphi h_\psi \mathbf{M} \Sigma - \gamma_b h_\pi \Sigma = 0,$$

where Σ is the variance-covariance matrix of the elements in \mathbf{z} . Postmultiplying this expression with Σ^{-1} and rearranging, we finally obtain

$$h_\pi [\mathbf{M} - \gamma_f \mathbf{M}^2 - \gamma_b \mathbf{I}] = \varphi h_\psi \mathbf{M},$$

which are exactly the cross-equation restrictions derived in Section 4. It is also immediately obvious from this development that GMM does not impose any uniqueness conditions.

References

- [1] Benhabib, J. and R.E.A. Farmer, "Indeterminacy and Sunspots in Macroeconomics," in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics 1A* (North-Holland: Elsevier Science B.V., 1998).
- [2] Bilts, M., "The Cyclical Behavior of Marginal Cost and Prices," *American Economic Review* 77 (1987), 838-857.
- [3] Bilts, M. and P.J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, forthcoming (October 2004).
- [4] Blanchard, O.J. and N. Kiyotaki, "Monopolistic Competition and the Effect of Aggregate Demand," *American Economic Review*, 77 (September 1987), 647-666.
- [5] Calvo, G. A., "Staggered Prices in an Utility-Maximizing Framework," *Journal of Monetary Economics*, 12 (1983), 383-398.
- [6] Clarida, R., J. Gali and M. Gertler, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* 37 (December 1999), 1661-1707.
- [7] Eichenbaum, M. and J.D.M. Fisher, "Evaluating the Calvo Model of Sticky Prices," mimeo, Northwestern University and Federal Reserve Bank of Chicago, June 2004.
- [8] Fuhrer, J.C. and G. Moore, "Inflation Persistence," *Quarterly Journal of Economics* 110 (February 1995), 127-153.
- [9] Fuhrer, J.C., G.R. Moore and S.D. Schuh, "Estimating the Linear Quadratic Inventory Model. Maximum Likelihood versus Generalized Method of Moments," *Journal of Monetary Economics* 110 (1995), 115-157.
- [10] Fuhrer, J.C., "The (Un)Importance of Forward-Looking Behavior in Price Specifications," *Journal of Money, Credit and Banking* 29 (August 1997), 338-350.
- [11] Fuhrer, J.C. and G. D. Rudebusch, "Estimating the Euler Equation for Output," *Journal of Monetary Economics* 51 (September 2004), 1133-1153.
- [12] Gali, J. and M. Gertler, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics* 44 (1999), 195-222.
- [13] Gali, J., M. Gertler, and J. D. Lopez-Salido, "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve," *Journal of Monetary Economics*, forthcoming (2003).

- [14] Goffe, W.L., "SIMANN: A Global Optimization Algorithm using Simulated Annealing," *Studies in Nonlinear Dynamics & Econometrics* 1 (October 1996), 169-176.
- [15] Hamilton, J. D., *Time Series Analysis* (Princeton: Princeton University Press, 1994).
- [16] Hansen, L.P. and T.J. Sargent, "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control* 2 (1980).
- [17] Ireland, P.N., "Sticky Price Models of the Business Cycle," *Journal of Monetary Economics* 47 (February 2001), 3-18.
- [18] Jondeau, E. and H. Le Bihan, "Testing for a Forward-Looking Phillips Curve. Additional Evidence from European and US Data," Erudite Doc. No. 01-05, Université Paris XII, August 2001.
- [19] Jondeau, E. and H. Le Bihan, "ML vs GMM Estimates of Hybrid Macroeconomic Models (With an Application to the New Phillips Curve)," Notes d'études et de recherche #103, Banque de France, October 2003.
- [20] Judd, K.L., *Numerical Methods in Economics* (Cambridge: MIT Press, 1998).
- [21] King, R.G. and M.W. Watson, "The Solution of Singular Linear Difference Systems Under Rational Expectations," *International Economic Review* 39 (November 1998), 1015-1026.
- [22] King, R.G. and A. Kurmann, "Expectations and the Term Structure of Interest Rates: Evidence and Implications," *Federal Reserve Bank of Richmond Economic Quarterly* 88 (Fall 2002), 49-95.
- [23] Kozicki, S. and P.A. Tinsley, "Alternative Sources of the Lag Dynamics of Inflation," mimeo, Federal Reserve of Kansas City and University of Cambridge, October 2002.
- [24] Lindé, J., "Estimating New-Keynesian Phillips Curves: A Full-Information Maximum Likelihood Approach," Working Paper Series No. 129, Sveriges Riksbank, December 2002.
- [25] Lubik, T. and F. Schorfheide, "Computing Sunspot Equilibria in Linear Rational Expectations Models," *Journal of Economic Dynamics and Control* 28 (2003), 273-285.
- [26] Ma, A., "GMM estimation of the new Phillips curve," *Economics Letters* 76 (2002), 411-417.
- [27] Nason, J.M. and G.W. Smith, "Identifying the New Keynesian Phillips Curve," mimeo, Federal Reserve Bank of Atlanta and Queens University May 2003.

- [28] Roberts, J.M., "Is Inflation Sticky?" *Journal of Monetary Economics* 39 (1997), 173-196.
- [29] Roberts, J.M., "How Well Does the New Keynesian Sticky-Price Model Fit the Data?" Finance and Economics Discussion Series 2001-13, Board of Governors of the Federal Reserve System, February 2001.
- [30] Roll, R., *The Behavior of Interest Rates: An Application of the Efficient Market Model to U.S. Treasury Bills* (New York: Basic Books, 1969).
- [31] Rotemberg, J.J. and M. Woodford, "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomic Annual* 12 (1997), 297-346.
- [32] Sargent, T.J., "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure." *Journal of Monetary Economics* 5 (1979), 133-143.
- [33] Sargent, T.J., *Macroeconomic Theory*, 2nd edition (San Diego: Academic Press, 1987).
- [34] Sbordone, A.M., "Prices and Unit Labor Costs: Testing Models of Pricing," *Journal of Monetary Economics* 49 (March 2002), 265-292.
- [35] Sbordone, A.M., "Do Expected Future Marginal Costs Drive Inflation Dynamics?" *Journal of Monetary Economics*, forthcoming (2004).
- [36] Smets, F. and R. Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1 (September 2003), 1123-1175.
- [37] Stock, J.H., J. Wright, and M. Yogo, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics* 20 (2002), 518-529.
- [38] Stock, J.H. and M. Yogo, "Testing for Weak Instruments in Linear IV Regression," mimeo, Harvard University, 2003.
- [39] Taylor, J.B. (1998), "Staggered Wage and Prices in Macroeconomics," in J.B. Taylor and M. Woodford, *Handbook of Macroeconomics*, North-Holland: Elsevier Science B.V.
- [40] Woodford, M., *Interest and Prices* (Princeton: Princeton University Press, 2003).
- [41] Wolman, A.L., "The Frequency and Costs of Individual Price Adjustment," *Federal Reserve of Richmond Economic Quarterly* 86 (Fall 2000), 1-22.

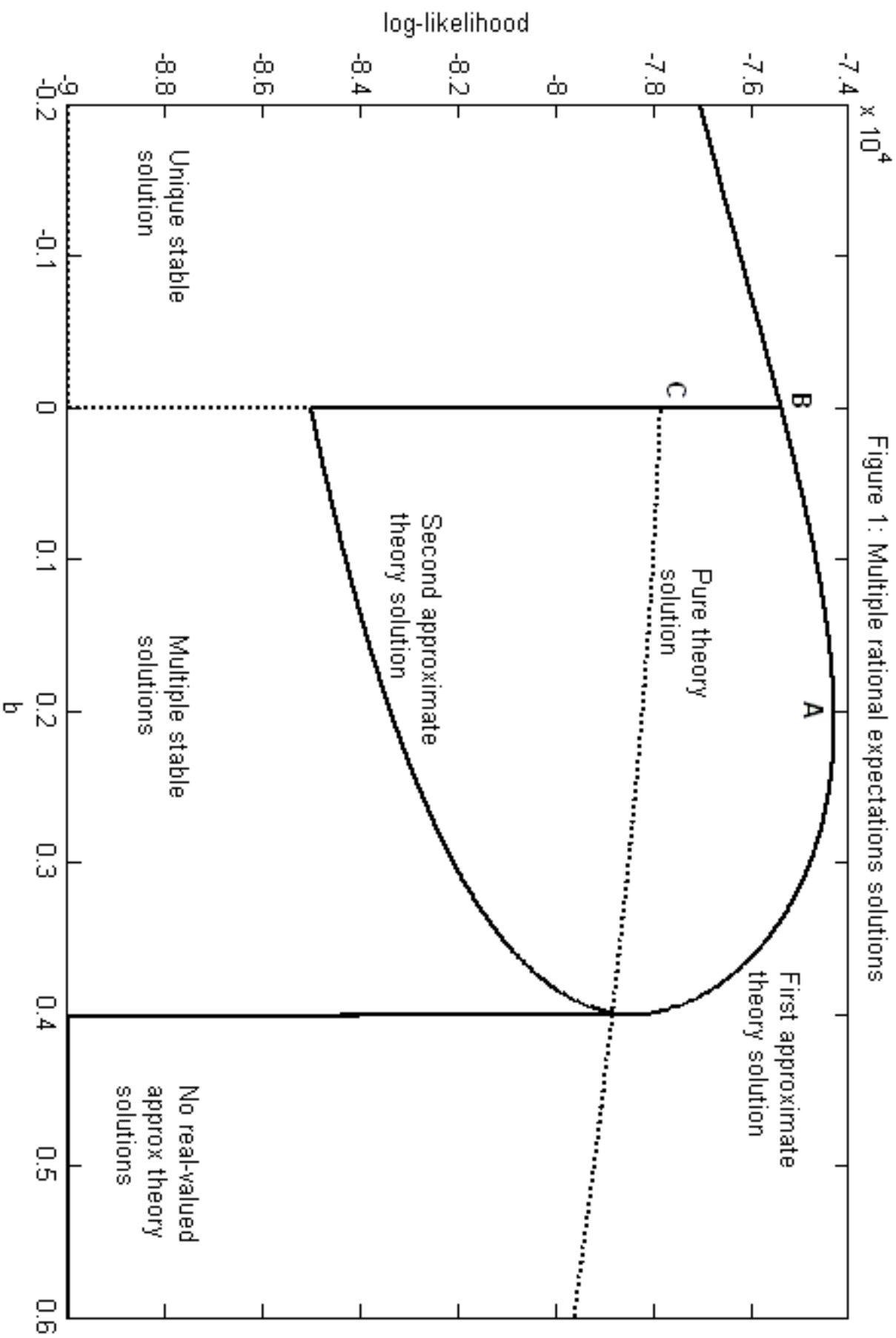


Table 1
Unrestricted VAR estimates

(a) Labor income share and inflation

	s_{t-1}	π_{t-1}	s_{t-2}	π_{t-2}	s_{t-3}	π_{t-3}	R^2
s_t	0.884 (0.083)	0.036 (0.244)	-0.004 (0.111)	-0.038 (0.289)	-0.078 (0.084)	0.271 (0.240)	0.780
π_t	0.074 (0.028)	0.642 (0.083)	-0.060 (0.038)	0.047 (0.099)	-0.014 (0.029)	0.240 (0.082)	0.824

(b) Output gap and inflation

	x_{t-1}	π_{t-1}	x_{t-2}	π_{t-2}	x_{t-3}	π_{t-3}	R^2
x_t	1.211 (0.081)	0.002 (0.003)	-0.173 (0.128)	-0.002 (0.004)	-0.092 (0.085)	-0.002 (0.003)	0.950
π_t	0.832 (1.969)	0.561 (0.081)	1.418 (3.113)	-0.012 (0.094)	0.635 (2.066)	0.235 (0.076)	0.840

Notes: This table reports the coefficient estimates for the OLS regressions of (a) U.S. labor income share and inflation on lags thereof, and (b) U.S. output gap and inflation on lags thereof. The sample period is 1961:1-1997:4. Standard errors are shown in brackets.

Table 2
Maximum Likelihood estimates

	Structural parameters			NKPC slope coefficients			Likelihood ratio	p-value
	β	θ	ω	γ_f	γ_b	φ		
<i>(a) Labor income share</i>								
Pure theory solution	1.000	0.923 (0.091)	0.000 (0.006)	1.000 (0.082)	0.000 (0.067)	0.006 (0.020)	29.097	0.000
Approximate theory solution	1.000	0.848 (0.046)	0.050 (0.029)	0.944 (0.147)	0.056 (0.146)	0.024 (0.015)	4.844	0.564
Uniqueness solution	1.000	0.874 (0.039)	0.740 (0.066)	0.542 (2.122)	0.458 (1.541)	0.003 (0.092)	22.417	0.001
<i>(b) Output gap</i>								
Pure theory solution	1.000	1.000 (502.040)	0.000 (0.100)	1.000 (0.075)	0.000 (0.060)	0.000 (0.692)	42.487	0.000
Approximate theory solution	1.000	0.538 (0.072)	0.437 (0.066)	0.552 (0.955)	0.448 (0.587)	0.123 (4.022)	19.526	0.003
Uniqueness solution	1.000	0.538 (0.072)	0.437 (0.066)	0.552 (0.955)	0.448 (0.587)	0.123 (4.022)	19.526	0.003

Notes: This table reports the ML estimates of the NKPC for the different solutions to the cross-equation restrictions. Estimates are reported for both the case where real marginal cost is measured by (a) U.S. labor income share, and (b) linearly detrended output gap. BHHH standard errors are reported in parenthesis.