THE UNIVERSITY OF SYDNEY

Economics Working Paper Series

2015 - 13

Comments on “Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve”

Fabiá Gumbau-Brisa & Denny Lie

May 2015

Comments on "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve"*

Fabià Gumbau-Brisa[†] Denny Lie[‡]

May 2015

Abstract

This comment reexamines the finding in Cogley and Sbordone ("Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review* 98(5): 2101-26, 2008) that the New Keynesian Phillips curve is purely forward-looking once time-varying trend inflation is accounted for. We perform various types of robustness analysis involving the second-stage estimation procedure, the number of indexation lags, and the vantage point of expectations. All in all, our estimates show that the main result in Cogley and Sbordone is not robust to these seemingly innocuous modifications.

JEL Classification: E12, E31, E52

Keywords: Cogley and Sbordone, closed form, model-consistent expectations, New Keynesian Phillips curve, forward-looking Euler equation, time-varying trend inflation,

*Fabià's work for this project was carried out in its entirety while employed by the Federal Reserve System. The views expressed in this paper are the authors' and do not necessarily reflect the views of any regional Federal Reserve Bank, the Federal Reserve System, or JPMorgan Chase.

[†]JPMorgan Chase. Email: fabia.gumbau-brisa@jpmorgan.com.

[‡]School of Economics, The University of Sydney. Email: denny.lie@sydney.edu.au.

1 Introduction

In a recent seminal paper ("Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review* 98(5): 2101-26, 2008) Timothy Cogley and Argia Sbordone show that the apparent need for a backward-looking inflation component in the New Keynesian Phillips Curve (NKPC) in order to fit the data can be done away with once the specification takes into account that, over the sample period considered, trend inflation was not constant. That is, once time-varying trend inflation — interpreted as the Federal Reserve's implicit inflation target — is explicitly factored into the estimation of the NKPC, the indexation parameter is estimated to be insignificant, and the NKPC is thus purely forward looking. Put another way, it is the persistence of the trend inflation, rather than an inherent feature of firms' price adjustment behavior, that imparts persistence to the inflation process. This result is important as the indexation assumption that generates the backward-looking component is not well-grounded in terms of its microeconomic foundation and is inconsistent with various evidence from micro price studies (e.g. Bils and Klenow, 2004 and Nakamura and Steinsson, 2008). Yet despite this important contribution, the robustness of this finding has not been sufficiently tested.

This paper reexamines the finding in Cogley and Sbordone (henceforth, CS) by performing various kinds of robustness analysis to ascertain whether their main finding still holds true under several important, but relatively minor, modifications.[1] In particular, our focus is on the modification to the estimation methodology involving the second stage. CS estimate the NKPC using a two-step procedure: the first step involves estimating an unrestricted VAR that can be used to proxy for expectations, while the second stage uses a minimum-distance estimator that exploits the cross-equation restrictions implied by the NKPC. As this is a partial-information method and the estimation sample is relatively small, the parameter estimates are potentially sensitive to the amount of model-consistent restrictions on expectations that is imposed in the second stage. We produce a number of estimates involving various degrees of model-consistent restrictions on expectations and compare our findings to the estimates in CS. Further to this, we perform other robustness analysis involving the number of indexation lags and the vantage point of expectations. We use the same data and sample period as in CS.

All in all, our findings show that the main result in CS has to be taken with some caution.

[1] A number of findings in this comment are based on the findings reported in our previous papers: Barnes, Gumbau-Brisa, Lie, and Olivei (2011) and Gumbau-Brisa, Lie, and Olivei (2011).

In particular, only in one case, the case considered in CS, we are able to obtain the result that the indexation parameter is effectively zero once time-varying trend inflation is accounted for. In all other cases, the indexation parameter is shown to be positive and economically significant at standard confidence levels. We refrain from completely ruling out time-varying trend inflation as an explanation to the inflation persistence puzzle as further analysis is needed in order to conclusively address this issue. Additional future work could possibly involve estimating the NKPC with time-varying trend using a full-information method and developing a criterion or test statistic to test whether the NKPC itself is misspecified in the first place.

2 Main comment: The role of model-consistent restrictions on expectations

2.1 The estimated NKPC in Cogley and Sbordone

CS consider a NKPC with time-varying trend inflation, derived from the Calvo staggered-price assumption. In addition, firms that do not get the opportunity to adjust prices optimally are assumed to index their prices to the previous lag of inflation, i.e.

$$P_t(i) = (\Pi_{t-1})^{\rho} P_{t-1}(i),$$

where $P_t(i)$ is the price set by firm i when it cannot reoptimize at time t and $\Pi_t = P_t/P_{t-1}$ is the period t gross rate of inflation. The parameter $\rho \in [0, 1]$ measures the degree of indexation to $t-1$ aggregate inflation. Trend inflation is assumed to be an exogenous process that evolves as a random walk process with drift. Log-linearizing the equations summarizing the firms' price adjustment decision around this time-varying trend then leads to a NKPC with time-varying coefficients and with additional elements involving higher-order terms for expected inflation, stochastic discount factor, and output growth. Denoting by hat the log-deviation of a variable from its long-run trend value, we can write the NKPC as follows

$$\widehat{\pi}_t = \rho(\widehat{\pi}_{t-1} - \widehat{g}_t^{\pi}) + \lambda_t E_t(\widehat{\pi}_{t+1} - \rho\widehat{\pi}_t) + \zeta_t \widehat{mc}_t + \gamma_t \widehat{D}_t + u_{\pi,t}, \tag{1}$$

where $\widehat{\pi}_t$ is net inflation at time t, $\widehat{mc}$ is the real marginal cost, $\widehat{g}_t^{\pi} = \ln(\overline{\Pi}_t/\overline{\Pi}_{t-1})$ is the growth rate of trend inflation, $u_{\pi,t}$ is a structural i.i.d. shock, and $\widehat{D_t}$ is defined recursively as

$$\widehat{D}_t = \varphi_{1,t} E_t(\widehat{q}_{t,t+1} + \widehat{g}_{t+1}^{y}) + \varphi_{1,t}(\theta - 1)E_t\left\{\widehat{\pi}_{t+1} - \rho\widehat{\pi}_t\right\} + \varphi_{1,t} E_t \widehat{D}_{t+1}. \tag{2}$$

$\widehat{q}_{t,t+1}$ is the real stochastic discount factor between time t and $t+1$ and $\widehat{g}_t^y$ is the time-t output growth rate. The time-varying coefficients, λ_t, ζ_t, γ_t, and $\varphi_{1,t}$, are functions of the structural parameters of the model.[2] Combining (1) and (2) yields

$$\begin{aligned}\widehat{\pi}_t &= \widetilde{\rho}_{1,t}^D\left(\widehat{\pi}_{t-1}-\widehat{g}_t^{\bar{\pi}}\right)+\widetilde{\zeta}_t^D\widehat{mc}_t+d_{1,t}^D E_t\widehat{\pi}_{t+1} \\ &\quad +d_{2,t}^D E_t\sum_{j=2}^{\infty}\varphi_{1,t}^{j-1}\widehat{\pi}_{t+j} \\ &\quad +d_{3,t}^D E_t\sum_{j=0}^{\infty}\varphi_{1,t}^{j}\left[\widehat{Q}_{t+j,t+j+1}+\widehat{g}_{t+j+1}^y\right] \ +\widetilde{u}_{\pi,t},\end{aligned} \tag{3}$$

where $\widehat{Q}_{t+j,t+j+1}=\widehat{q}_{t+j,t+j+1}+\widehat{\pi}_{t+j+1}$ is the nominal stochastic discount factor between time $t+j$ and $t+j+1$. The coefficient $\varphi_{1,t}$ in (2) is bounded at the estimation stage so that the infinite sum is finite.[3] This is the form of the NKPC estimated in CS.[4] The main interest is on the estimate of the coefficient attached to the backward-looking inflation, $\widetilde{\rho}_{1,t}^D$ — if it is zero, the NKPC is purely forward looking. Since $\widetilde{\rho}_{1,t}^D\equiv\rho/\Delta_t$, $\widetilde{\rho}_{1,t}^D=0$ if the indexation parameter ρ is estimated to be zero.[5]

2.2 The estimation procedure in CS

CS use a two-stage procedure to estimate the structural parameters in (3).[6] The first stage involves estimating a Bayesian VAR with time-varying trends,

$$\mathbf{z}_t=\boldsymbol{\mu}_t+\mathbf{A}_t\mathbf{z}_{t-1}+\boldsymbol{\varepsilon}_{z,t}, \tag{4}$$

where $\boldsymbol{\varepsilon}_{z,t}$ is a possibly heteroskedastic but serially uncorrelated error vector and the coefficients in $\boldsymbol{\mu}_t$ and $\mathbf{A}_t$ are assumed to evolve as random walk. The evolution of the coefficients in $\mathbf{A}_t$ is constrained so that the roots of $\mathbf{A}_t$ at each point in time lie inside the unit circle. $\mathbf{z}_t$ is the vector of variables that includes output growth, real marginal cost (unit labor cost), inflation, and the Fed funds rate. For any variable y_t in $\mathbf{z}_t$, we can express its conditional expectation using the VAR above as

$$E_t\widehat{y}_{t+k}=\mathbf{e}_y'\mathbf{A}_t^k\widehat{\mathbf{z}}_t, \tag{5}$$

[2] See Appendix A in Cogley and Sbordone (2008) for definitions on the time-varying coefficients.

[3] This condition is equivalent to equations (39) and (40) in Cogley and Sbordone (2008) being satisfied. These two equations are imposed in the estimation stage.

[4] We use slightly different notations to CS in regards to the time-varying coefficients. See Appendix A for details on the coefficients.

[5] $\Delta_t=1+\rho\lambda_t+\gamma_t(\theta-1)\rho\varphi_{1,t}>0$.

[6] As mentioned in CS (2008), although ideally we would like to use a full-information method in a Bayesian setting such a method is computationally intractable for the problem in hand.

where $\widehat{\mathbf{z}}_t$ is the vector of variables expressed in deviations from the time-varying trends

$$\widehat{\mathbf{z}}_t \equiv \mathbf{z}_t - (\mathbf{I} - \mathbf{A}_t)^{-1}\boldsymbol{\mu}_t,$$

and $\mathbf{e}'_y$ is the selection vector for variable y_t in $\mathbf{z}_t$.

The second stage involves a minimum distance estimator, exploiting the cross-equation restrictions implied by the NKPC and the first-stage VAR estimates. These cross-equation restrictions can be obtained using the forecasting rule in (5), given the posterior distribution produced in the first stage VAR. Based on information at $t-1$ the conditional expectation of inflation based on the NKPC (3) is given by

$$\begin{aligned}
\mathbf{e}'_\pi \mathbf{A}_{t-1}\widehat{\mathbf{z}}_{t-1} &= \widetilde{\rho}^D_{1,t-1}\mathbf{e}'_\pi \mathbf{I}\widehat{\mathbf{z}}_{t-1} + \widetilde{\zeta}^D_{t-1}\mathbf{e}'_{mc}\mathbf{A}_{t-1}\widehat{\mathbf{z}}_{t-1} \\
&\quad + d^D_{1,t-1}\mathbf{e}'_\pi \mathbf{A}^2_{t-1}\widehat{\mathbf{z}}_{t-1} + d^D_{2,t-1}\varphi_{1,t-1}\mathbf{e}'_\pi \mathbf{J}_{t-1}\mathbf{A}^3_{t-1}\widehat{\mathbf{z}}_{t-1} \\
&\quad + d^D_{3,t-1}(\mathbf{e}'_Q \mathbf{J}_{t-1}\mathbf{A}_{t-1}\widehat{\mathbf{z}}_{t-1} + \mathbf{e}'_{g^y}\mathbf{J}_{t-1}\mathbf{A}^2_{t-1}\widehat{\mathbf{z}}_{t-1}),
\end{aligned}$$

where $\mathbf{J}_t \equiv (\mathbf{I} - \varphi_{1,t}\mathbf{A}_t)^{-1}$.[7] Dropping $\widehat{\mathbf{z}}_{t-1}$ leads to the following cross-equation restrictions

$$\begin{aligned}
\mathbf{e}'_\pi \mathbf{A}_{t-1} &= \widetilde{\rho}^D_{1,t-1}\mathbf{e}'_\pi \mathbf{I} + \widetilde{\zeta}^D_{t-1}\mathbf{e}'_{mc}\mathbf{A}_{t-1} \\
&\quad + d^D_{1,t-1}\mathbf{e}'_\pi \mathbf{A}^2_{t-1} + d^D_{2,t-1}\varphi_{1,t-1}\mathbf{e}'_\pi \mathbf{J}_{t-1}\mathbf{A}^3_{t-1} \\
&\quad + d^D_{3,t-1}(\mathbf{e}'_Q \mathbf{J}_{t-1}\mathbf{A}_{t-1} + \mathbf{e}'_{g^y}\mathbf{J}_{t-1}\mathbf{A}^2_{t-1}) \qquad (6) \\
&\equiv \mathbf{g}^D(\boldsymbol{\mu}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\psi}),
\end{aligned}$$

where $\boldsymbol{\psi}$ denotes the vector of structural parameters to be estimated. The relevant distance for the estimation is then

$$\mathbf{F}^D_1(\boldsymbol{\mu}_t, \mathbf{A}_t, \boldsymbol{\psi}) \equiv \mathbf{e}'_\pi \mathbf{A}_t - \mathbf{g}^D(\boldsymbol{\mu}_t, \mathbf{A}_t, \boldsymbol{\psi}) = \underline{0}', \qquad (7)$$

which must for all periods t. Equation (7) says that if the NKPC in (3) is the true data generating process for inflation, the reduced form forecast ($\mathbf{e}'_\pi \mathbf{A}_t$) and the NKPC-based forecast for inflation must be the same. Since $\widehat{\mathbf{A}}_t$ is estimated with sampling error this relationship is not exactly satisfied in practice. The second stage procedure is then akin to finding the structural parameter estimates so that the distance between the two forecasts is as close as possible to zero, given the posterior distribution of $\widehat{\mathbf{A}}_t$ estimated in the first stage.

[7] Note that this relationship is valid under the condition that $|\varphi_{1,t}| < 1$, which is imposed at the estimation stage (equations (39) and (40) in Cogley and Sbordone).

As shown in CS, when trend inflation varies over time there is an additional long-run relationship that needs to be imposed in the estimation stage, given by

$$F_2(\mu_t, A_t, \psi) \equiv \left(1 - \alpha \overline{\Pi}_t^{(1-\rho)(\theta-1)}\right)^{(1+\theta\omega)/(1-\theta)} \left[\frac{1 - \alpha \overline{q} \overline{g}^y \overline{\Pi}_t^{\theta(1+\omega)(1-\rho)}}{1 - \alpha \overline{q} \overline{g}^y \overline{\Pi}_t^{(\theta-1)(1-\rho)}}\right] - (1-\alpha)^{(1+\theta\omega)/(1-\theta)} \frac{\theta}{\theta - 1} \overline{mc}_t = \underline{0}'. \tag{8}$$

where $\overline{y}_t$ denotes the trend value of variable y_t at time t. On the structural parameters, α is the Calvo probability of price fixity, θ is the elasticity of substitution across goods, and ω is the elasticity of firm's marginal cost with respect to its own output (a measure of strategic complementarity in price setting). To allow for identification ω (and the discount factor, β) is not estimated — hence, $\psi = \{\rho, \alpha, \theta\}$.[8] The complete set of cross-equation restrictions that need to be satisfied in any given period t is then given by

$$\begin{aligned} \mathcal{F}^D(\mathbf{\Theta}) &\equiv \left[\mathcal{F}_1^D, ..., \mathcal{F}_T^D\right] \\ \mathcal{F}_t^D &\equiv \left[\mathbf{F}_1^D(\boldsymbol{\mu}_t, \mathbf{A}_t, \boldsymbol{\psi}),\ \mathbf{F}_2(\boldsymbol{\mu}_t, \mathbf{A}_t, \boldsymbol{\psi})\right], \end{aligned}$$

where

$$\mathbf{\Theta} \equiv \{\boldsymbol{\mu}_t, \mathbf{A}_t\}_{t=1}^T.$$

The parameter estimates for each ensemble $i \in [1, M]$ (from the set of posterior distribution obtained in the first stage) can then be obtained from

$$\widehat{\boldsymbol{\psi}}_i^D \equiv \arg\min \mathcal{F}^D(\widehat{\mathbf{\Theta}}_i) \cdot \mathcal{F}^D(\widehat{\mathbf{\Theta}}_i)' \quad \text{for } i = 1, .., M. \tag{9}$$

2.3 The role of model-consistent restrictions on expectations in the second stage

We first note that the estimation procedure above is a partial-information method in which expectations in the second stage are proxied using an unrestricted VAR estimated in the first stage. These expectations are therefore *not* necessarily model-consistent. A natural question of interest is then on how much model discipline on expectations should be imposed in the second stage and in particular, whether the estimates in CS change once we add additional model-consistent restrictions on expectations.

[8] CS set $\beta = 0.99$ and $\omega = 0.43$. We use these values for all the cases considered.

Consider the estimation procedure used in CS described above. The second-stage procedure used in CS, summarized in (9), uses cross-equation restrictions based on the NKPC in (3), which is essentially a difference equation. The expectations of forward-looking variables in (3) are then directly proxied using the forecasting rule (5) given the first-stage VAR estimates, producing the cross-equation restrictions (6). Hence, one can say that the procedure used in CS essentially imposes *no* model-consistent restriction on expectations, beyond those used to derive the NKPC in (3). We can, however, impose additional model discipline on expectations at the estimation stage. In the context of the NKPC considered in CS, there are two ways to do this: *partial* and *complete* (*closed-form*) impositions. Below we first briefly discuss these two cases separately before providing the link between them in the subsequent discussion.

2.3.1 Partial imposition of model-consistent expectations

One way to impose additional model discipline on expectations is by solving forward for one period the expected inflation in (1). Define $\widehat{\pi}_t^n \equiv \widehat{\pi}_t - \rho(\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}})$ as the portion of inflation that is not predetermined at time t so that we can write equation (1) as

$$\widehat{\pi}_t^n = \lambda_t E_t \widehat{\pi}_{t+1}^n + \zeta_t \widehat{mc}_t + \gamma_t \widehat{D}_t + u_{\pi,t}$$

For $E_t \widehat{\pi}_{t+1}^n$ to be model-consistent it must be the case that

$$E_t \widehat{\pi}_{t+1}^n = \lambda_t E_t \widehat{\pi}_{t+2}^n + \zeta_t E_t \widehat{mc}_{t+1} + \gamma_t E_t \widehat{D}_{t+1}$$

Substituting this expression for $E_t \widehat{\pi}_{t+1}^n$ in the previous equation leads to

$$\begin{aligned} \widehat{\pi}_t^n &\equiv \widehat{\pi}_t - \rho(\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}}) \\ &= \lambda_t^2 E_t \widehat{\pi}_{t+2}^n + \zeta_t \left(\widehat{mc}_t + \lambda_t E_t \widehat{mc}_{t+1}\right) + \gamma_t \left(\widehat{D}_t + \lambda_t E_t \widehat{D}_{t+1}\right) + u_{\pi,t} \\ &= \lambda_t^2 E_t \left(\widehat{\pi}_{t+2} - \rho \widehat{\pi}_{t+1}\right) + \zeta_t \left(\widehat{mc}_t + \lambda_t E_t \widehat{mc}_{t+1}\right) + \gamma_t \left(\widehat{D}_t + \lambda_t E_t \widehat{D}_{t+1}\right) + u_{\pi,t} \end{aligned}$$

Combining the above equation with (2) in a similar manner when deriving (3) leads to a new form of the NKPC in which we impose model discipline on expected inflation for one more period. At this point we note that although we impose only partial restrictions on inflation expectations through (1) instead of the the full NKPC equation in (3), these restrictions are nevertheless model-consistent and thus is a valid NKPC to be estimated in a similar way to the second stage estimation of (3).

Next, proxying for current and expected variables as before, one can show that the cross-equation restrictions based on this new NKPC is related to (6) and (7) by (arguments of the functions are removed for ease of exposition)

$$\mathbf{F}_{1,t}^{D(1)} = \mathbf{F}_{1,t}^{D} \cdot (\mathbf{I} + \lambda_t \mathbf{A}_t) \tag{10}$$

where the supercripts $D(1)$ indicates that here we impose model-consistent restrictions on expectations for one more period compared to the base case estimated in CS. The complete set of cross-equation restrictions that need to be satisfied in any given period t is then given by

$$\begin{aligned} \mathcal{F}^{D(1)}(\mathbf{\Theta}) &\equiv \left[\mathcal{F}_1^{D(1)}, ..., \mathcal{F}_T^{D(1)}\right] \\ \mathcal{F}_t^{\ D(1)} &\equiv \left[\mathbf{F}_{1,t}^{\ D(1)},\ \mathbf{F}_{2,t}\right], \end{aligned}$$

so that the second stage procedure becomes

$$\widehat{\boldsymbol{\psi}}^{D(1)} \equiv \arg\min_{\boldsymbol{\psi}} \mathcal{F}^{\ D(1)}(\widehat{\mathbf{\Theta}}_i) \cdot \mathcal{F}^{\ D(1)}(\widehat{\mathbf{\Theta}}_i))', \tag{11}$$

Notice that the long run restrictions $\mathbf{F}_2\left(\boldsymbol{\mu}_t, \mathbf{A}_t, \boldsymbol{\psi}\right)$ and the first-stage VAR estimates $\widehat{\mathbf{A}}_t$ remain unchanged. The only difference between the estimation problems (9) and (11) is that (11) uses more information about the model (NKPC) dynamics — and this additional information is not exploited in (9). In the estimation problem (9), the only information used to characterize expectations is the unrestricted linear projections obtained through $\widehat{\mathbf{A}}_t$, which does not take explicitly into account that the behavior of future inflation should also satisfy the NKPC relationship.

Of course, instead of just one more period, one can also further discipline expectations for any additional periods $j > 0$. This involves solving the expectations of inflation in the RHS of (3) for j periods forward, producing cross-equation restrictions

$$\mathbf{F}_{1,t}^{D(j)} \equiv \mathbf{F}_{1,t}^{D} \cdot (\mathbf{I} + \lambda_t \mathbf{A}_t + ... + \lambda_t^j \mathbf{A}_t^j)$$

The higher j is the higher is the degree of model-consistent restrictions on expectations that we impose at the estimation stage. At the limit, one can also solve forward for inflation expectations ad infinitum ($j \longrightarrow \infty$), yielding

$$\begin{aligned}\widehat{\pi}_t &= \rho(\widehat{\pi}_{t-1} - \widehat{g}_t^{\pi}) + \zeta_t \sum_{i=0}^{\infty} \lambda_t^i E_t \widehat{mc}_{t+i} + \gamma_t \sum_{i=0}^{\infty} \lambda_t^i E_t \widehat{D}_{t+i} + u_{\pi,t} \qquad (12)\\ &= \rho(\widehat{\pi}_{t-1} - \widehat{g}_t^{\pi}) + \zeta_t \sum_{j=0}^{\infty} \lambda_t^j E_t mc_{t+i}\\ &\quad +\gamma_t \varphi_{1,t} \sum_{j=0}^{\infty} \lambda_t^j \sum_{i=0}^{\infty} \varphi_{1,t}^i E_t \left\{ q_{t+i+j,t+1+i+j} + g_{t+1+i+j}^y + (\theta - 1)\pi_{t+1+i+j}^n \right\} + u_{\pi,t}\end{aligned}$$

We leave the derivation of the above NKPC with its cross-equation restrictions in Appendix B. For this ad-infinitum case it can be shown that the cross-equation restrictions are related to $\mathbf{F}_{1,t}^D$ by

$$\mathbf{F}_{1,t}^{D(\infty)} = \mathbf{F}_{1,t}^D \cdot (\mathbf{I} - \lambda_t \mathbf{A})^{-1}$$

and hence, the minimum distance problem in the second stage becomes

$$\widehat{\boldsymbol{\psi}}^{D(\infty)} \equiv \arg\min_{\boldsymbol{\psi}} \mathcal{F}^{\,D(\infty)}(\widehat{\boldsymbol{\Theta}}_i) \cdot \mathcal{F}^{\,D(\infty)}(\widehat{\boldsymbol{\Theta}}_i))', \qquad (13)$$

where

$$\begin{aligned}\mathcal{F}^{D(\infty)}(\boldsymbol{\Theta}) &\equiv \left[\mathcal{F}_1^{D(\infty)}, ..., \mathcal{F}_T^{D(\infty)}\right]\\ \mathcal{F}_t^{\,D(\infty)} &\equiv \left[\mathbf{F}_{1,t}^{\,D(\infty)},\ \mathbf{F}_{2,t}\right],\end{aligned}$$

Notice that the second stage problem in (13) contains the same number of equations (restrictions) as in (9) and (11). This means that the scale of the estimation problem is the same irrespective of the degree of model-consistent restrictions on expectations being imposed.

At this point we note that the NKPC in (12) is not yet in *closed form* as expected inflation terms still appear on the right-hand side. Still, compared to the DE specification considered in CS, the estimation problem in (13) contains additional model-consistent restrictions on expectations. Furthermore when $\gamma_t \to 0$ equation (12) becomes a closed-form NKPC where we have solved out all expected inflation terms and inflation becomes a function of its driving process only (current and expected real marginal costs). In fact, the estimates of γ_t tends to hover around zero for all ensembles in all the cases we consider below. We also note that to derive (12), or more appropriately, to estimate this NKPC using the minimum distance estimator in (13), we require the roots of $\lambda_t \mathbf{A}_t$ to lie inside the unit circle, i.e. $\|\lambda_t \mathbf{A}_t\| < 1$.[9] If this condition is not satisfied

[9] A more stringent condition is $|\lambda_t| < 1$. However, for the estimation strategy used in Cogley and Sbordone, the condition $\|\lambda_t \mathbf{A}_t\| < 1$ is sufficient.

the infinite sums (12) will diverge and the estimation problem in (13) is not well defined. This condition is not guaranteed by model conditions alone, and is therefore an empirical issue. We defer discussing this issue, along with the consequences if $\|\lambda_t \mathbf{A}_t\| \geq 1$, until the next subsection.

Table 1: Structural parameter estimates (median and 90% trust region)
Imposing various degree of model-consistent restrictions on expectations
Sample period: 1960.Q1–2003.Q4

	ρ	α	θ
DE	0 (0.00,0.16)	0.583 (0.45,0.67)	9.76 (7.69,12.46)
DE(1)	0.28 (0.01,0.66)	0.627 (0.50,0.72)	11.44 (9.45,16.23)
DE(2)	0.29 (0.00,0.69)	0.650 (0.52,0.75)	10.71 (8.76,15.14)
DE(4)	0.49 (0.19,0.89)	0.667 (0.47,0.80)	11.55 (9.46,32.63)
DE(∞)	0.61 (0.29,0.90)	0.838 (0.73,0.91)	10.05 (7.91,14.57)
CF	0.70 (0.41,0.91)	0.873 (0.77,0.93)	9.88 (7.88,13.29)

Notes: (1) numbers in parentheses are 90% trust regions; (2) DE(j) refers to DE estimates imposing j additional periods of model-consistent restrictions on expectations; (3) CF refers to the exact closed-form estimates where we solve for all expected inflation terms prior to estimating the parameters in the second stage estimation procedure.

Table 1 displays the estimates of the structural parameters when we impose various degrees of model-consistent restrictions on expectations discussed above. In the table *DE* refers to the estimation case considered in CS, summarized in (9), while *DE(j)* refers to the case where we impose model discipline on expectations for j additional period. As can be seen from the table we replicate the main result in CS that the indexation parameter ρ is zero once time-varying trends are taken into account—the NKPC is therefore purely forward looking. Estimates of the Calvo parameter, α, and the elasticity of substitution, θ, are also virtually identical. The estimate of ρ, however, starts to turn positive once additional model consistency is imposed. Even when we only impose additional model consistency for one additional period (*DE(1)*), the median estimate of ρ jump to 0.28, which indicates that the backward-looking inflation component is still present. While the 90% trust region in *DE(1)* (and *DE(2)*) still includes 0, imposing model-consistent restrictions on expectations for four periods (*DE(4)*) already points to significantly different conclusion than the *DE* case. Here $\rho = 0.49$ and the 90% trust region doesn't include 0. $\rho = 0.49$ means that the coefficient on past inflation in the NKPC is about 0.32 on average, which is smaller than the coefficient on the forward-looking inflation but is not quantitatively negligible. Finally, at the limit, imposing model consistency for infinite periods (*DE(∞)*) the median estimate of the indexation parameter is further away from zero ($\rho = 0.61$) and zero is *not* in the 90% trust region.

The estimates of α and θ are also sensitive to the degree of model-consistent restrictions on expectations being imposed. For example for *DE(∞)*, the median estimate of α jumps to 0.838 from 0.583 estimated in CS. This implies that price re-optimization now occurs every 5-6 quarters on average. The median estimate of θ is now about 10, which implies a steady-state markup of 11.1% when inflation is zero. In all, the results in Table 1 indicate that the estimates of the structural parameters in the NKPC with time-varying trend are sensitive to how much model-consistent restrictions on expectations being imposed, at least in the partial-information estimation procedure considered in CS.

2.3.2 Complete imposition of model-consistent expectations (closed-form)

As noted above even though in *DE(∞)* we impose model-consistent restrictions on expectations for infinite periods, the resulting NKPC is still not entirely closed form. It is possible to derive an exact closed-form (CF) version of the NKPC by also solving forward the expected inflation terms contained in (2). The resulting expression is

$$\begin{aligned}\widehat{\pi}_t &= \rho\left(\widehat{\pi}_{t-1}-\widehat{g}_t^{\pi}\right)+\zeta_t\sum_{j=0}^{\infty}\xi_{1,t}^{j}\sum_{i=0}^{\infty}\xi_{2,t}^{i}E_t\left\{mc_{t+i+j}-\varphi_{1,t}mc_{t+1+i+j}\right\} \\ &\quad +\gamma_t\varphi_{1,t}\sum_{j=0}^{\infty}\xi_{1,t}^{j}\sum_{i=0}^{\infty}\xi_{2,t}^{i}E_t\left\{q_{t+i+j,t+1+i+j}+g_{t+1+i+j}^{y}\right\}+u_{\pi,t},\end{aligned} \tag{14}$$

where $\xi_{1,t}+\xi_{2,t}=\lambda_t+\varphi_{1,t}+\gamma_t\varphi_{1,t}(\theta-1)$ and $\xi_{1,t}\xi_{2,t}=\lambda_t\varphi_{1,t}$. We leave the derivation of the closed-form NKPC above, the definition of the coefficients, and the corresponding cross-equation restrictions in Appendix C.[10] Using the cross-equation restrictions based on (14) in the second stage estimation problem is akin to imposing the highest possible degree of model-consistent restrictions

[10] The closed-form NKPC in (14) is also derived and discussed in Barnes, Gumbau-Brisa, Lie, and Olivei (2011). In their paper, the specification (14) is referred as the "exact closed form."

on expectations. This is the complete opposite to the restrictions based on (6), which essentially imposes no model consistency at the estimation stage.

Note that the parameters $\xi_{1,t}$ and $\xi_{2,t}$ need to be bounded in order to guarantee that the geometric sums in (14) are well defined. For the second-stage estimation strategy considered here, we require the roots of $\{\xi_{i,t}\mathbf{A}_t\}_{i=1,2}$ to lie inside the unit circle, i.e. $\delta_t^{\max} \equiv \max\{\|\xi_{i,t}\mathbf{A}_t\|\}_{i=1,2} < 1$, which is again, an empirical issue. We return to this issue in the next subsection, but we mention here that this condition is also related to the uniqueness of the closed form NKPC in (14) and the validity of the reduced-form VAR in (4). Note that when $\gamma_t \to 0$ we have $\xi_{1,t} \to \lambda_t$ and $\xi_{2,t} \to \varphi_{1,t}$, and as a result the closed form (14) and the *DE(*∞) representation in (12) are the same.[11] As long as γ_t is small, these two representations of the NKPC are very similar — this condition is generally satisfied for all the cases we consider here, including for the specification considered in CS.

The last row of Table 1 shows the estimates based on the closed-form (CF) NKPC in (14). Once again, the median estimate of ρ is significantly away from zero and the 90% trust region doesn't include 0. The estimate of $\rho = 0.71$ means that the coefficient on lagged inflation is closer to 0.5 than 0, e.g. as assumed in Christiano, Eichenbaum, and Evans (2005) under the constant trend assumption. And not surprisingly, since γ_t tends to be small, all the parameter estimates in *CF* are close to the estimates obtained in *DE(*∞). In contrast to case considered in CS, imposing complete consistency on model dynamics point to the presence of lagged inflation in NKPC even under time-varying trend inflation. The CF estimates also suggest that when all of the model restrictions on expectations are taken into account, the link from marginal costs to inflation becomes weaker.

2.4 Discussion

2.4.1 General discussion

A natural question to ask from the estimation results in Table 1 is on the conditions under which the estimates are the same irrespective of the degree of model-consistent restrictions on expectations being imposed. One condition that this occurs is when the system of implicit equations $\mathbf{F}_{1,t}^{D(j)}$ is just identified, i.e. the number of equations is equal to the number of parameter values to be

[11] To see this point, notice that

$$\sum_{j=0}^{\infty}\lambda_t^j\sum_{i=0}^{\infty}\varphi_{1,t}^i E_t\left\{mc_{t+i+j}-\varphi_{1,t}mc_{t+1+i+j}\right\}=\sum_{j=0}^{\infty}\lambda_t^j E_t mc_{t+j}.$$

.

estimated. This case obviously rarely occurs in estimations involving time-series macroeconomics data. A more relevant case is when the system is overidentified in which the number of parameters to be estimated is far less than the number of equations. In this case $\widehat{\boldsymbol{\psi}}^{D(j)}$ will be the same for all j only when the true matrix $\mathbf{A}$ is known. When $\mathbf{A}$ is known or estimated precisely, then it does not matter which specification of the NKPC is being estimated, since inflation forecasts generated from the reduced-form VAR are perfectly model-consistent — that is, $\mathbf{F}^{D}_{1,t}$, and hence $\mathbf{F}^{D(j)}_{1,t}$ $\forall$ j, are exactly equal zero. But since in practice this matrix is bound to be estimated with sampling errors, the parameter estimates are going to be different for different j. This is true even when the NKPC is the true data generating process for inflation and $\widehat{\mathbf{A}}$ is the same in all j.

The discussion above raises another important question: Which second-stage estimation strategy should we choose when the system is overidentified and $\mathbf{A}$ is estimated with sampling errors? The answer depends on whether the NKPC provides a good characterization of inflation dynamics in the data. As long as the NKPC is close to the true data generating process for inflation, imposing additional model discipline on expectations will always yield parameter estimates that are closer to their true values. Additional model information being imposed at the estimation stage is always good as long as the information itself is consistent with the data. Things may be different, however, if the NKPC is a poor approximation to the inflation process. In this case, being agnostic about the way expectations are formed may provide better, less-biased estimates — hence, imposing no model discipline on expectations as in CS may be the better option. Deciding which way to go is beyond the scope of this comment and thus, we leave this issue for future research. At the minimum, we need to design a specification test to test whether the NKPC in (3) is misspecified in the first place to conclusively answer this question. We note, however, that our analysis below on the issue of determinacy seems to better support the estimates in *DE(*∞) and *CF*, compared to those in Cogley and Sbordone.

2.4.2 Determinacy and the validity of parameter estimates

To derive the closed-form specification in (14) the parameters $\xi_{1,t}$ and $\xi_{1,t}$ need to be bounded in order to guarantee the geometric sums are well defined. In particular, for the cross-equation restrictions to be well defined the condition

$$\delta_t^{\max} \equiv \max\{\left\|\xi_{i,t}\mathbf{A}_t\right\|\}_{i=1,2} < 1 \tag{15}$$

must be satisfied. Here we perform various checks to see whether this condition is satisfied based on the resulting structural parameter estimates.

It is important to note that this condition must also be satisfied regardless of which specification is being estimated — this includes the specification (3) considered in Cogley and Sbordone. As discussed in Appendix D, if this condition does not hold, the NKPC does not have a unique forward solution and hence, is incompatible with any structural model with a reduced form VAR representation as in (4), unless one is willing to make additional assumptions on other structural equations besides the NKPC. This condition is thus a necessary and sufficient condition for the determinacy of the NKPC solution. In the presence of indeterminacy, the reduced-form VAR would instead involve an infinite number of lags of the endogenous variables, or equivalently the error term would have a moving average component.[12] Neither of these two representation would be consistent with the law of motion in (4). Then, trying to match the cross-equation restrictions of the NKPC to a reduced form VAR representation (4) that the model cannot possibly admit would invalidate the estimation procedure. In essence, the estimation procedure adopted in Cogley and Sbordone assumes a reduced form as in (4), and hence implicitly assumes that $\delta_t^{\max} < 1$. Empirical violations

[12] Even when an infinite-order VAR representation can be well approximated by a finite-order VAR, it is still necessary for the condition $\delta_t^{\max} \equiv \max\{\left\|\xi_{i,t}\mathbf{A}_t\right\|\}_{i=1,2} < 1$ to be satisfied. This is because if this condition does not hold, the forecasting rule in (5) is misspecified. To illustrate this point, consider for example the case in which $\|\xi_1\mathbf{A}\| > 1 > \|\xi_2\mathbf{A}\|$. Now the closed form can be written as

$$\begin{aligned}\pi_t^n &= \frac{1}{\xi_{1,t-1}}\pi_{t-1}^n - \frac{\zeta_{t-1}}{\xi_{1,t-1}}\sum_{i=0}^{\infty}\xi_{2,t-1}^i E_{t-1}\left\{mc_{t-1+i} - \varphi_{t-1}mc_{t+i}\right\} \\ &\quad -\frac{\gamma_{t-1}\varphi_{t-1}}{\xi_{1,t-1}}\sum_{i=0}^{\infty}\xi_{2,t-1}^i E_{t-1}\left\{q_{t-1+i,t+i} + g_{t+i}^y\right\} \\ &\quad -\frac{1}{\xi_{1,t-1}}u_{\pi,t-1} + \eta_{\pi,t},\end{aligned}$$

where η_π is an expectational error, that is,

$$\eta_{\pi,t} \equiv \pi_t^n - E_{t-1}\left\{\pi_t^n\right\}.$$

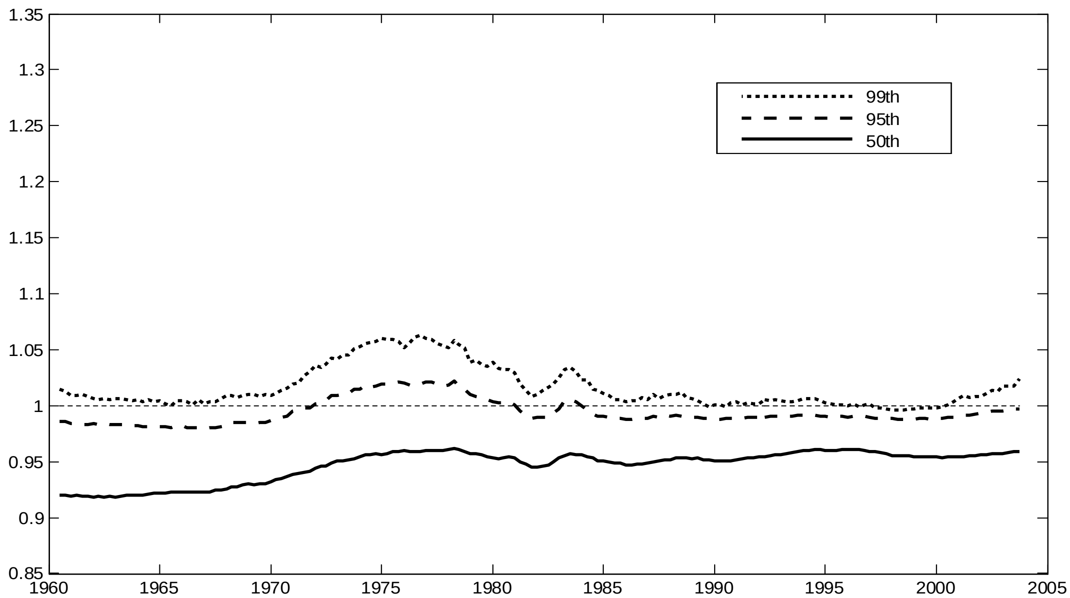
The solution above includes the time $t-1$ structural shock for inflation, $u_{\pi,t-1}$. Unlike the CF solution (14), this lagged structural shock is not guaranteed to cancel out with any of the elements in the driving process. Indeed, the CF solution in (14), which is the unique determinate solution, is the only solution that does not involve predictable error terms. Instead, the solution above is consistent with a general equilibrium model that is not invertible, and hence, has a moving average error term. When the data-generating process for inflation follows the indeterminate solution above, forecasts obtained from (5) are incorrect because they are omitting relevant information. Under the above, the correct forecasting rule is

$$\begin{aligned}E_{t-j}\left\{\pi_{t+1}\right\} &= \mathbf{e}_\pi'\mathbf{A}_{t-j}^{j+1}\mathbf{z}_{t-j} - \mathbf{e}_\pi'\mathbf{A}_{t-j}^{j}\mathbf{b}_{t-j}u_{\pi,t-j} \\ &\neq \mathbf{e}_\pi'\mathbf{A}_{t-j}^{j+1}\mathbf{z}_{t-j},\end{aligned}$$

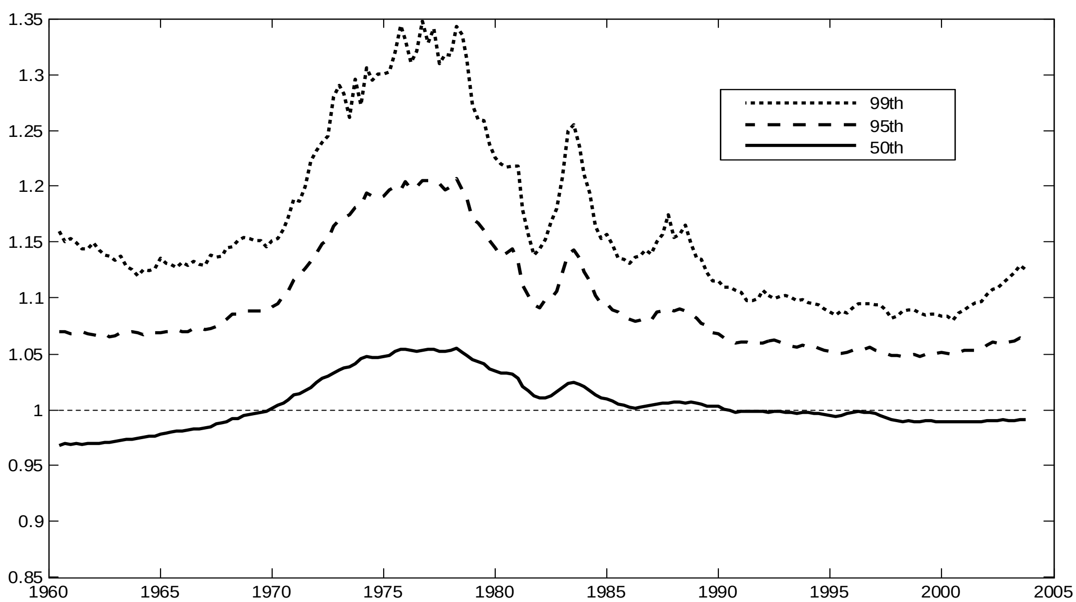
where the vector $\mathbf{b} \neq \mathbf{0}$ relates the reduced-form shocks (the vector $\boldsymbol{\varepsilon}$), to the structural shock u_π. This same issue arises when $\|\xi_i\mathbf{A}\| > 1$ for $i = 1, 2$

Figure 1: Distribution of largest root of $\delta_t^{\max}$ in absolute value.
Median, 95th and 99th percentiles.
(Same $\hat{\mathbf{A}}_t$ estimates in both panels)

Panel A: Closed-Form (CF), distribution of $\delta_t^{\max} \equiv \max\left\{\left\|\xi_{i,t}\mathbf{A}_t\right\|\right\}_{i=1,2}$.



Panel B: Difference-Equation (DE) Form, distribution of $\delta_t^{\max} \equiv \max\left\{\left\|\xi_{i,t}\mathbf{A}_t\right\|\right\}_{i=1,2}$



of this condition would render meaningless the estimates for *all* the specifications we consider.[13]

Panel A in Figure 1 depicts the distribution of $\delta_t^{\max}$ for the *CF* specification in (14). Violations of the condition $\delta_t^{\max} < 1$ are relatively infrequent. The 95th percentile of the distribution is always below unity with the exception of a few violation during the 1970s. In all, the figure illustrates that the conditions for the validity of the CF estimates are generally satisfied. The corresponding findings for the *DE* specification in (3) considered in Cogley and Sbordone are sharply different, indicating an incompatibility between the first- and second-stage estimates. In this case, reported in panel B of Figure 1, the distribution of $\delta_t^{\max}$ always has the 95th percentile above unity. Moreover, while the median (50th percentile) of the distribution in *CF* is always below unity, there is a considerable number of periods in which the median of the distribution in *DE* fails to satisfy the determinacy condition in (15). Worse yet, the fraction of ensembles for which the *DE* estimates never violate the condition (15) over the estimated sample period is a mere 0.17 percent. In contrast, using this more restrictive criterion the corresponding fraction for the *CF* estimates is 30.44 percent.

We also note that, as discussed in Appendix C,

$$\max\left\{\xi_{1,t},\ \xi_{2,t}\right\} \geq \lambda_t \ .$$

It is clear then the condition $\|\lambda_t \mathbf{A}_t\| < 1$ discussed in section 2.3.1 is just a necessary (weaker) condition for determinacy and hence, for the existence of the closed form in (14). As long as the condition (15) is satisfied, we also have $\|\lambda_t \mathbf{A}_t\| < 1$. Yet, as mentioned previously, $\max\left\{\xi_{1,t},\ \xi_{2,t}\right\} = \lambda_t$ if $\gamma_t = 0$. As it turns out, the estimates of γ_t are negligibly small and thus, the condition $\|\lambda_t \mathbf{A}_t\| < 1$ is a good approximation to the necessary and sufficient condition for determinacy (15).[14] An implication of this is that the *DE(∞)* specification in (12) is a good approximation to the closed-form (*CF*) specification in (14) — not surprisingly, the parameter estimates based on the two specifications are not far off (see the last two rows of Table 1). The implication in terms of the distribution of $\|\lambda_t \mathbf{A}_t\|$ for the *DE* and *CF* specifications are also similar to that in Figure 1.[15] Using this weaker determinacy condition, the fraction of ensembles for which the *DE* and *CF* estimates never violate the condition $\|\lambda_t \mathbf{A}_t\| < 1$ over the estimated sample period is 10.70 percent and 68.41 percent, respectively.

---

[13] This is an issue that arises in limited information estimation, when estimates for $\mathbf{A}_t$ and $\{\xi_{i,t}\}_{i=1,2}$ are obtained at different stages.

[14] For the closed-form (*CF*) estimates with indexation to two lags of inflation (unconstrained τ) we obtain that $\left|\gamma_t \varphi_{1,t} (\theta - 1)\right|$ has a median of $8.7x10^{-4}$, a mode of $2.8x10^{-6}$, and is below $3x10^{-3}$ for all ensembles.

[15] Results are available upon request.

The discussion above raises the issue of whether the estimates in Table 1 change once we remove ensembles for which the determinacy condition (15) does not hold. Table 2 reports the estimates for the *CF* specification for the ensembles that do not violate the necessary and sufficient condition for determinacy. The first row of Table 2 reports the estimates for the *CF* specification irrespective of whether the determinacy condition is satisfied, first reported in the last row of Table 1. The other rows of Table 2 report median estimates for ensembles that do not violate the condition (15) for more than 100-, 95-, and 90-percent of the time, respectively. The most restrictive criterion (100 percent) allows to preserve only those ensembles for which no violations ever occur over the 174 quarters. It is apparent from the table that regardless of the criterion chosen, the estimates are remarkably similar. This implies that for the *CF* specification, the distribution of parameter estimates is insensitive to whether the determinacy condition (15) is satisfied.

Table 2: Structural parameter estimates (median and 90% trust region)
Preserving ensembles that satisfy $\delta_t^{\max} < 1$
Sample period: 1960.Q1–2003.Q4

	ρ	α	θ
CF	0.70 (0.41,0.91)	0.873 (0.77,0.93)	9.88 (7.88,13.29)
CF (100% criterion)	0.67 (0.41,0.89)	0.867 (0.78,0.92)	9.98 (7.86,12.42)
CF (95% criterion)	0.71 (0.43,0.90)	0.874 (0.79,0.93)	9.87 (7.94,12.79)
CF (90% criterion)	0.71 (0.43,0.91)	0.874 (0.79,0.93)	9.89 (7.90,13.11)

Notes: (1) numbers in parentheses are 90% trust regions; (2) The x% criterion preserves ensembles that satisfy the determinacy condition for x% of the 174 quarters; (3) The number of optimized ensembles is 4271; (4) The 100% criterion retains 30.4% of the ensembles, the 95% retains 78.7%, and the 90% retains 92.2%.

Table 3 reports the same exercise of preserving only ensembles that satisfy the determinacy condition for the *DE* and *DE(j)* specifications in Table 1, for the 100% criterion. For these cases, we use the weaker determinacy condition, $\|\lambda_t \mathbf{A}_t\| < 1$ — similar results are obtained when the condition (15) is used instead. Comparing the estimates in Table 3 to those in Table 1, it is clear that the distribution of estimates is also largely unaltered even when we remove "indeterminate ensembles". In particular, it is still the case that only in the specification considered in Cogley and

Sbordone—the *DE* case in the first row—that we are able to conclude that the median estimate of the indexation parameter ρ is zero.

It is also possible to estimate the *DE* specification in a manner that guarantees determinacy and hence, compatibility between the two estimation stages. In the exercise reported in the first row of Table 4, we explicitly impose the condition that $\delta_t^{\max} < 1$ (for all t) in the second-stage estimation. The median estimate of ρ is now much higher, although it is less precisely estimated. In fact, the stark similarity of the median estimates of ρ between the *DE* case in Table 4 ($\rho = 0.71$) and the *CF* case in Table 1 ($\rho = 0.70$) seems to indicate that the finding in Cogley and Sbordone that inflation is purely forward-looking after accounting for time-varying trends comes from biased estimates due to the presence of indeterminacy.

Table 3: Structural parameter estimates (median and 90% trust region)
Table 1 *DE* estimates after removing ensembles with $\|\lambda_t \mathbf{A}_t\| \geq 1$ (100% criterion)
Sample period: 1960.Q1–2003.Q4

	ρ	α	θ
DE	0	0.618	10.37
	(0.00,0.09)	(0.45,0.69)	(8.91,12.55)
DE(1)	0.26	0.657	11.82
	(0.01,0.63)	(0.54,0.76)	(10.04,14.76)
DE(2)	0.28	0.674	11.22
	(0,0.72)	(0.53,0.79)	(9.18,14.49)
DE(4)	0.52	0.682	11.64
	(0.20,0.93)	(0.54,0.84)	(9.98,15.07)
DE(∞)	0.55	0.830	10.07
	(0.26,0.85)	(0.73,0.90)	(7.92,13.55)

Notes: (1) numbers in parentheses are 90% trust regions; (2) DE retains 10.70%, DE(1) retains 15.34%, DE(2) retains 18.83%, DE(4) retains 27.83%, DE(∞) retains 59.32% of the ensembles.

Finally, one could argue that assuming $\delta_t^{\max} < 1$ in order to derive the closed form NKPC in (14) may bias the estimates towards satisfying this condition. We have shown in Table 1, however, that the estimates $\widehat{\boldsymbol{\psi}}^{D(4)}$ (*DE(4)*), for example, are already close to the CF estimates. As the *DE(4)* specification does not rely on the condition $\delta_t^{\max} < 1$ for its derivation, this seems to indicate that this concern is not an issue. Not surprisingly then, the distribution of $\delta_t^{\max}$ for $\widehat{\boldsymbol{\psi}}^{D(4)}$ (not shown) is still largely below unity. Moreover, further increasing the number of model consistent restrictions on expectations shifts this distribution towards the one based on the *CF* estimates shown in panel A of Figure 1.[16]

Table 4: *DE* estimates
Restricted to determinate solution ($\delta_t^{\max} < 1$, $\forall t$)
Sample period: 1960.Q1–2003.Q4

	ρ	α	θ	τ
DE	0.71	0.590	7.96	1
	(0.00, 0.93)	(0.34, 0.83)	(5.69, 10.92)	–
DE_ uncon	0.85	0.608	11.47	0.60
	(0.57,1.00)	(0.54,0.76)	(10.09,13.34)	(0.42,0.81)

Notes: (1) point estimates are medians; (2) numbers in parentheses are 90% trust regions; (3) *DE* refers to the difference-equation specification considered in Cogley and Sbordone with indexation to one lag of inflation ; (4) *DE_ uncon* refers to the difference-specification but with indexation to a weighted average of two lags of inflation, with τ represents the weight given to $t-1$ inflation relative to $t-2$ inflation.

Another evidence toward repelling this concern comes from a particular case considered in Cogley and Sbordone that omits terms involving the discount factor, output growth, and terms involving higher-order leads of inflation — the estimates for this case are reported in Table C.3 in their paper.[17] These terms do not appear to be empirically relevant for our estimates (due to small estimates of γ_t). For this set of estimates, we also shut down time variation in the NKPC coefficients.[18] Note that in this case $\delta_t^{\max}$ is simply equal to β (the discount rate), and as a result, the stability requirement that $\|\beta \mathbf{A}_t\| < 1$ is always satisfied. In practical terms, this means that the second-stage estimates cannot produce evidence contrary to the implicit assumption in the first stage of model determinacy.[19] Estimation results are reported in Table 5 for the *DE* and *CF* specifications (the first row and the third row, respectively). It is evident that the findings are largely unaltered from a qualitative standpoint. Quantitatively, the changes in the estimates are small across the two specifications. This indicates that the differences between the *DE* and *CF*

[16] Results for $\widehat{\boldsymbol{\psi}}^{D(j)}$, where $j = \{2, 4, 6, 8\}$, and are available upon request.

[17] As mentioned in Cogley and Sbordone, this case is equivalent to assuming that firms who are unable to adjust prices optimally fully index their prices to a mixture of current trend inflation and one-period lagged inflation.

[18] The estimation procedure still uses the time-varying VAR in (4) to form expectations, and the long-run restriction (8).

[19] Even with a standard NKPC with constant coefficients it is possible to have indeterminate equilibria - for instance if policy does not satisfy the Taylor principle (see Lubik and Schorfheide 2003). In our exercise this potential outcome is ruled out by assumption, since the VAR is assumed to properly characterize the unique stationary dynamics of the model, and because $\beta < 1$ (the real discount factor) is not estimated.

estimates do not depend on whether the condition $\delta_t^{\max} < 1$ is met, and are instead the consequence of the CF specification exploiting additional model-consistent restrictions.

Table 5: Structural parameter estimates(median and 90% trust region)
Constant coefficients and removing higher-order leads
Sample period: 1960.Q1–2003.Q4

	ρ	α	θ	τ
DE	0	0.562	12.08	1
	(0.00,0.11)	(0.44,0.66)	(8.22,15.22)	—
DE_uncon	0.70	0.612	11.72	0.58
	(0.37,1.00)	(0.52,0.73)	(10.21,13.95)	(0.32,0.81)
CF	0.42	0.777	11.62	1
	(0.17,0.71)	(0.68,0.85)	(8.47,15.10)	—
CF_uncon	0.87	0.864	12.29	0.68
	(0.72,0.99)	(0.78,0.92)	(10.77,16.06)	(0.48,0.89)

Notes: (1) numbers in parentheses are 90% trust regions; (2) *DE* and *CF* correspond to difference-equation (DE) and closed-form (CF) specifications with $\tau = 1$ (indexation to one lag of inflation); (3) *DE_uncon* and *CF_uncon* correspond to DE and CF cases with unconstrained τ (indexation to two lags of inflation), respectively.

3 Other comments and issues

In this section we touch upon two other robustness issues. The first one involves the structure of the indexation of non-optimized prices: instead of one lag of inflation we now allow for two lags of inflation in the indexation mechanism. This modification, while technically minor, is important from an empirical standpoint as it can reduce the effect of misspecification bias on the estimates. The indexation mechanism now takes the form

$$P_t(i) = (\Pi_{t-1}^{\tau}\Pi_{t-2}^{1-\tau})^{\rho} P_{t-1}(i)$$

As previously, the parameter $\rho \in [0, 1]$ measures the degree of indexation, while $\tau \in [0, 1]$ represents the weight given to $t - 1$ aggregate inflation relative to $t - 2$ aggregate inflation. This indexation mechanism nests the one-lag inflation indexation case when $\tau = 1$. We leave the full details of the derivation of the resulting NKPC, including the cross-equation restrictions needed for the second-stage estimation, to Appendix A.

Table 6 replicates the estimates in Table 1 under this new indexation mechanism. It is apparent that allowing for the second lag of indexation dramatically alters the structural parameter estimates for all specifications. In particular, the median estimates of the indexation parameter ρ are now much higher — this is true even for the difference-equation specification (DE) considered in Cogley and Sbordone ($\rho = 0.64$). Moreover, the 90% trust region does not include zero. The value of $\rho = 0.64$ implies that the coefficient on the backward-looking inflation terms in the NKPC is roughly 0.4, a sharp departure from the finding in Cogley and Sbordone. The same conclusion applies to the closed-form specification (*CF)*: the median estimate of ρ increases from 0.70 in the one-lag indexation case to 0.89 in the present case. The value of $\rho = 0.89$ implies the backward-looking and the forward-looking inflation terms receive approximately the same weight in the NKPC. In terms of the estimates of τ, the median estimates are above 0.5 for all specifications, suggesting that the first lag of inflation is relatively more important in the indexation mechanism. Moreover, the 90% trust regions of τ are bounded well away from 1 for all specifications.

Table 6: Structural parameter estimates (median and 90% trust region)
(With indexation to two lags of indexation)
Sample period: 1960.Q1–2003.Q4

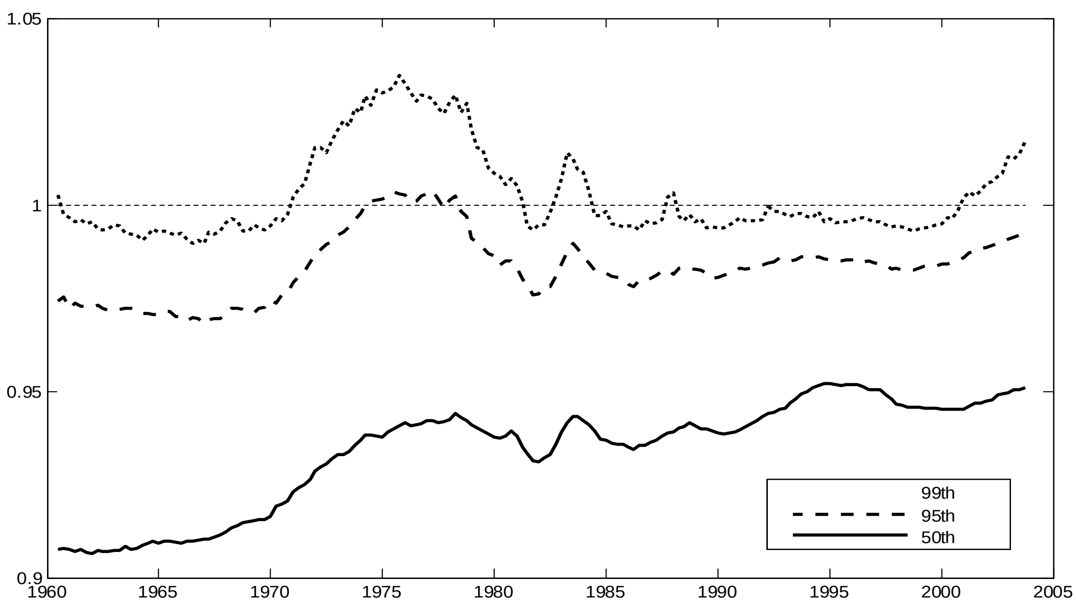
	ρ	α	θ	τ
DE	0.64 (0.29,1.00)	0.597 (0.51,0.72)	11.87 (10.25,14.79)	0.56 (0.23,0.80)
DE(1)	0.75 (0.44,1.00)	0.646 (0.56,0.74)	11.82 (10.28,14.74)	0.72 (0.51,0.95)
DE(2)	0.81 (0.50,1.00)	0.673 (0.59,0.77)	11.78 (10.30,15.00)	0.68 (0.45,0.93)
DE(4)	0.86 (0.59,1.00)	0.722 (0.64,0.81)	11.75 (10.33,14.89)	0.71 (0.50,0.96)
$DE(\infty)$	0.89 (0.76,1.00)	0.877 (0.79,0.93)	12.27 (10.76,19.36)	0.69 (0.49,0.90)
CF	0.90 (0.80,0.98)	0.884 (0.82,0.93)	12.38 (10.77,20.81)	0.69 (0.50,0.90)

Notes: (1) numbers in parentheses are 90% trust regions; (2) DE(j) refers to DE estimates imposing j additional periods of model-consistent restrictions on expectations; (3) CF refers to the exact closed-form estimates where we solve for all expected inflation terms prior to estimating the parameters in the second stage estimation procedure; (4) τ is the weight given to $t-1$ inflation relative to $t-2$ inflation.

In terms of determinacy of the NKPC solution and the validity of parameter estimates, Figure 2 reports the distribution of $\delta_t^{\max}$ for the CF specification with two lags of indexation. As is apparent from Figure 2, allowing for two lags of indexation improves the validity of parameter estimates and makes the two estimation stages to be more compatible. Violations of the condition $\delta_t^{\max} < 1$ are now more infrequent, compared to the corresponding one-lag indexation case in panel A of Figure 1 — for example, except for a short period in mid-1970s, the 95th percentile of the distribution is well below one. For this *CF* specification the fraction of ensembles for which the estimates never violate the condition $\delta_t^{\max} < 1$ over the estimated sample period is now 59.91%.[20] The corresponding fraction for the *DE* specification with two lags of indexation (figure not shown), however, is only 10.78%, which is an improvement compared to the one-lag indexation case (a mere 0.17%). These numbers indicate again that the two-stage estimation strategy seems to suit the *CF* specification better than the *DE* specification.

Figure 2: Distribution of largest root of $\delta_t^{\max}$ in absolute value.
Median, 95th and 99th percentiles.
Closed-Form (CF) NKPC, indexation to two lags of inflation

Distribution of $\delta_t^{\max} \equiv \max\left\{\left\|\xi_{i,t}\mathbf{A}_t\right\|\right\}_{i=1,2}$.



We also report the estimates for the *DE* specification when we explicitly impose the condition that $\delta_t^{\max} < 1$. As shown in the second row of Table 4, allowing for indexation to two lags of inflation makes the parameter estimates to be closer to those in *CF* specification — this is especially true

[20] Using the weaker determinacy condition $\|\lambda_t \mathbf{A}_t\| \geq 1$, the fraction jumps to 80.78%.

for the indexation parameter ρ. When we consider the particular specification that omits terms involving the discount factor, output growth, and terms involving higher-order leads of inflation (Table 5), a similar conclusion to the one-lag indexation case is obtained. For both difference-equation and closed-form specifications (DE_uncon and CF_uncon in Table 5), the estimates are qualitatively unaltered compared to their corresponding cases in Table 6. This again indicates that the differences between the DE and CF estimates do not depend on whether the condition $\delta_t^{\max} < 1$ is met. And not surprisingly, when we remove ensembles that violates the determinacy condition (akin to Table 3 in the one-lag indexation assumption), the estimates are also largely unaltered.

Table 7: Changing the vantage point of expectations
DE estimates with one-lag of indexation

	ρ	α	θ
$t-1$ expectations (CS)	0 (0.00,0.16)	0.583 (0.45,0.67)	9.76 (7.69,12.46)
$t-2$ expectations	0.15 (0.00,0.68)	0.575 (0.49,0.66)	11.86 (9.67,15.99)
$t-3$ expectations	0.54 (0.04,0.94)	0.609 (0.51,0.72)	11.81 (10.25,14.09)
$t-4$ expectations	0.65 (0.21,0.94)	0.631 (0.54,0.72)	11.89 (10.31,14.33)
$t-5$ expectations	0.76 (0.44,0.96)	0.664 (0.57,0.74)	11.71 (10.27,13.79)
$t-6$ expectations	0.76 (0.44,0.95)	0.664 (0.57,0.74)	11.72 (10.27,13.88)
$t-7$ expectations	0.77 (0.47,0.95)	0.672 (0.58,0.74)	11.71 (10.27,13.85)
$t-8$ expectations	0.79 (0.46,0.97)	0.670 (0.58,0.74)	11.68 (10.27,13.86)

Notes: (1) the estimates in the first row is the case considered in Cogley and Sbordone where the expectations are taken at time $t-1$ in deriving the cross-equation restrictions; (2) subsequent specifications differ only in the time perspective at which expectations are taken in deriving the cross-equation restrictions.

Finally, we perform robustness check on the point in time from which expectations are taken.[21] Table 7 displays the estimates for the DE specification with indexation to one lag of inflation considered in Cogley and Sbordone. The first row of Table 7 refers to the vantage point of expectations

[21] More specifically, here we refer to the timing of expectations applied on both sides of (3) in order to obtain the cross-equation restrictions used in the second stage of the estimation. The intuition behind this is that if

$$E_{t-j}\{\pi_t\} = E_{t-j}\{NKPC_t\}$$

holds for $j = 1$, then it should hold for $j = 2, 3, ...$, as a result of the Law of Iterated Expectations.

used in Cogley and Sbordone ($t-1$ expectations). As shown in the remaining rows, the parameter estimates do not seem to be robust to the time perspective at which expectations are taken in deriving the cross-equation restrictions. The table shows that taking expectations further back in time increases the (median) weight given to lagged inflation, from zero in the case considered in Cogley and Sbordone to 0.65 when expectations are taken from the perspective of $t-4$. This lack of robustness does not seem to arise for the estimates using the closed-form (CF) specification.[22]

4 Conclusion

This paper reexamines the finding in Cogley and Sbordone (2008) that the New Keynesian Phillips curve (NKPC) is purely forward-looking once we account for time-varying trend inflation. We perform various robustness analysis involving the second-stage estimation procedure, the number of indexation lags, and the vantage point of expectations.

All in all our analysis shows that the main result in Cogley and Sbordone should be taken with some caution, at the very least for the sample period that they are considering. In particular, only in one specification — the specification considered in Cogley and Sbordone — we are able to obtain the result that the indexation parameter is zero once time-varying trend inflation is accounted for. In all other cases considered in this comment, the indexation parameter is shown to be positive and economically significant at standard confidence levels. We refrain from completely ruling out time-varying trend inflation as an explanation to the inflation persistence puzzle, as further analysis is needed to conclusively address this issue. This possibly involves estimating the NKPC with time-varying trend using a full-information method and developing a criterion or test statistic to test whether the NKPC itself is misspecified in the first place. We leave this issue for future research.

[22] Results are available upon request.

References

[1] Barnes, M.L., F. Gumbau-Brisa, D. Lie & G.P. Olivei, 2011. "Estimation of Forward-Looking Relationships in Closed Form: An Application to the New Keynesian Phillips Curve," Working Papers 11-3, Federal Reserve Bank of Boston.

[2] Bils, M. and P.J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112(5): 947–985.

[3] Calvo, G., 1983. "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3), pp. 383–398.

[4] Christiano, L., M. Eichenbaum, and C. Evans, 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", *Journal of Political Economy*, Vol. 113, pp. 1–45.

[5] Cogley, T. and A.M. Sbordone, 2008. "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review,* 98(5), pp. 2101–2126.

[6] Fernandez-Villaverde, J., Rubio-Ramirez, J., Sargent, T., and M. Watson, 2007. "A, B,C, (and D's) for Understanding VARs," *American Economic Review*, 97(3), pp. 1021-1026.

[7] Gumbau-Brisa, F., D. Lie & G.P. Olivei, 2011. "A response to Cogley and Sbordone's comment on "Closed-Form Estimates of the New Keynesian Phillips Curve with Time-Varying Trend Inflation"," Working Papers 11-4, Federal Reserve Bank of Boston.

[8] Kreps, D., 1998. "Anticipated Utility and Dynamic Choice," 1997 Schwartz Lecture, in *Frontiers of Research in Economic Theory*, D.P. Jacobs, E. Kalai, and M. Kamien (eds.), Cambridge University Press.

[9] Lubik, T. and F. Schorfheide, 2003. "Computing Sunspot Equilibria in Linear Rational Expectations Models," *Journal of Economic Dynamics and Control* 28, pp. 273-285.

[10] Lubik, T. and F. Schorfheide, 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), pp. 190–217.

[11] Nakamura, E. and J. Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models." *Quarterly Journal of Economics* 123(4): 1415–1464.

[12] Pesaran, M. H., 1989. *The Limits to Rational Expectations.* Oxford, Basil Blackwell.

[13] Sims, C. A. and T. Zha, 2006. "Does Monetary Policy Generate Recessions?" *Macroeconomic Dynamics* 10, pp. 231-27.

[14] Sims, E. R, 2009. "Non-Invertibilities and Structural VARs," University of Notre Dame, working paper.

"Comments on 'Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve'" (Gumbau-Brisa and Lie, 2015)

APPENDICES

[For online publication only.]

May 2015

Appendix A: The definition of coefficients in the NKPC with time-varying trends in difference equation (DE) form

In this appendix, we derive the NKPC in difference-equation (DE) form estimated in Cogley and Sbordone (2008) – the derivations steps closely follows their paper. In Cogley and Sbordone inflation is indexed to only one lag $(t-1)$ of inflation. Here we generalize the indexation mechanism so that the indexation is based on a mixture of the previous two lags of inflation, with weight τ and $(1-\tau)$ given to $t-1$ and $t-2$ inflation, respectively. The case considered in Cogley and Sbordone is a special case where $\tau = 1$. The parameter ρ is still interpreted as the degree of indexation to past inflation even in this more general case.

First, we derive the log-linear approximation of the evolution of aggregate prices. Let X_t be the optimal nominal price at time t chosen by firms that are allowed to adjust their prices, which happens at constant per-period probability $(1-\alpha)$. Based on our indexation mechanism, the price of an individual firm i that is not allowed to adjust (with probability α) evolves according to

$$P_t(i) = (\Pi_{t-1}^{\tau}\Pi_{t-2}^{1-\tau})^{\rho} P_{t-1}(i) \ .$$

Hence, the aggregate price based on the CES aggregator is given by

$$P_t = \left[(1-\alpha)X_t^{1-\theta} + \alpha\left\{(\Pi_{t-1}^{\tau}\Pi_{t-2}^{1-\tau})^{\rho} P_{t-1}\right\}^{1-\theta}\right]^{\frac{1}{1-\theta}} \ .$$

Dividing by the price level P_t, we have

$$1 = (1-\alpha)x_t^{1-\theta} + \alpha\left\{(\Pi_{t-1}^{\tau}\Pi_{t-2}^{1-\tau})^{\rho}\Pi_t^{-1}\right\}^{1-\theta} \ , \qquad \text{(A1)}$$

where x_t is the optimal relative price at time t. Next define stationary variables $\widetilde{\Pi}_t = \Pi_t/\overline{\Pi}_t$, $g_t^{\bar{\pi}} = \overline{\Pi}_t/\overline{\Pi}_{t-1}$, $g_t^{y} = Y_t/Y_{t-1}$, and $\widetilde{x}_t = x_t/\overline{x}_t$. Here, for any variable k_t, $\overline{k}_t$ is its time-varying trend. Equation (A1) can then be transformed in terms of these stationary variables to yield (after some algebra):

$$
\begin{aligned}
1 &= (1-\alpha)\widetilde{x}_t^{1-\theta}\overline{x}_t^{1-\theta} \\
&+\alpha\left[\begin{array}{c} \widetilde{\Pi}_{t-2}^{\rho(1-\tau)(1-\theta)}\widetilde{\Pi}_{t-1}^{\rho\tau(1-\theta)}\widetilde{\Pi}_t^{-(1-\theta)}\overline{\Pi}_t^{(1-\rho)(\theta-1)} \cdot \\ (g_{t-1}^{\bar{\pi}})^{-\rho(1-\tau)(1-\theta)}(g_t^{\bar{\pi}})^{-\rho(1-\tau)(1-\theta)}(g_t^{\bar{\pi}})^{-\rho\tau(1-\theta)} \end{array}\right] .
\end{aligned} \tag{A2}
$$

In the steady state where $\widetilde{x}_t = \widetilde{\Pi}_t = g_t^{\bar{\pi}} = 1$, (A2) can be solved for $\overline{x}_t$ as a function of $\overline{\Pi}_t$:

$$
\overline{x}_t = \left[\frac{1-\alpha\overline{\Pi}_t^{(1-\rho)(\theta-1)}}{1-\alpha}\right]^{\frac{1}{1-\theta}} . \tag{A3}
$$

Defining $\widehat{\pi}_t \equiv \ln\widetilde{\Pi}_t \equiv \ln(\Pi_t/\overline{\Pi}_t)$ and $\widehat{x}_t \equiv \ln\widetilde{x}_t$, imposing (A3), and rearranging, the log-linear approximation of (A2) around the steady state can be expressed as

$$
\begin{aligned}
\widehat{x}_t &= -\frac{1}{\varphi_{0,t}}\rho(1-\tau)\left(\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\bar{\pi}} - \widehat{g}_t^{\bar{\pi}}\right) \\
&\quad -\frac{1}{\varphi_{0,t}}\rho\tau\left(\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}}\right) \\
&\quad +\frac{1}{\varphi_{0,t}}\widehat{\pi}_t ,
\end{aligned} \tag{A3}
$$

where $\varphi_{0,t} = \frac{1-\alpha\overline{\Pi}_t^{(1-\rho)(\theta-1)}}{\alpha\overline{\Pi}_t^{(1-\rho)(\theta-1)}}$.

Next, we take the log-linear approximation to the first-order condition (FOC) of firms' pricing problem. Identically to the one-lag indexation case in Cogley and Sbordone (2008), the firms' FOC can be written as

$$
E_t\sum_{j=0}^{\infty}\alpha^j Q_{t,t+j}Y_{t+j}P_{t+j}\Psi_{tj}^{1-\theta}\left(X_t^{(1+\theta\omega)} - \frac{\theta}{\theta-1}MC_{t+j}\Psi_{tj}^{-(1+\theta\omega)}P_{t+j}^{\theta\omega}\right) = 0 , \tag{A4}
$$

where $Q_{t,t+j}$ and MC_{t+j} are the nominal discount factor and average marginal cost at $t+j$, respectively. The variable Ψ_{tj} enters in the CES demand function for any good i, $Y_{t+j}(i) = Y_{t+j}\left(\frac{P_{t+j}(i)\Psi_{tj}}{P_{t+j}}\right)$, with

$$
\Psi_{tj} = \begin{cases} 1 & j = 0 \\ \prod_{k=0}^{j-1}\left(\Pi_{t+k}^{\tau}\Pi_{t+k-1}^{1-\tau}\right)^{\rho} & j \geq 1 \end{cases} \tag{A5}
$$

The second line of (A5) makes clear that prices are indexed to a weighted average of the first two lags of inflation if they are not set optimally. Combining (A4) and (A5) and rearranging leads to

$$
X_t^{1+\theta\omega} = \frac{C_t}{D_t} ,
$$

where C_t and D_t are recursively defined by

$$\begin{aligned} C_t &= \frac{\theta}{\theta-1} Y_t P_t^{\theta(1+\omega)-1} MC_t \\ &\quad + E_t \left[\alpha q_{t,t+1} \Pi_t^{-\rho\tau\theta(1+\omega)} \Pi_{t-1}^{-\rho(1-\tau)\theta(1+\omega)} C_{t+1} \right] \end{aligned} \tag{A6}$$

$$\begin{aligned} D_t &= Y_t P_t^{\theta-1} \\ &\quad + E_t \left[\alpha q_{t,t+1} \Pi_t^{\rho\tau(1-\theta)} \Pi_{t-1}^{\rho(1-\tau)(1-\theta)} D_{t+1} \right] , \end{aligned} \tag{A7}$$

where $q_{t,t+1}$ now is the real discount factor. Defining the stationary variables $\widetilde{C}_t = \frac{C_t}{Y_t P_t^{\theta(1+\omega)}}$ and $\widetilde{D}_t = \frac{D_t}{Y_t P_t^{\theta-1}}$, we have based on (A6) and (A7):

$$\begin{aligned} \widetilde{C}_t &= \frac{\theta}{\theta-1} mc_t \\ &\quad + E_t \left[\alpha q_{t,t+1} g_{t+1}^y \Pi_{t+1}^{\theta(1+\omega)} \Pi_t^{-\rho\tau\theta(1+\omega)} \Pi_{t-1}^{-\rho(1-\tau)\theta(1+\omega)} \widetilde{C}_{t+1} \right] \end{aligned} \tag{A8}$$

$$\widetilde{D}_t = 1 + E_t \left[\alpha q_{t,t+1} g_{t+1}^y \Pi_{t+1}^{(\theta-1)} \Pi_t^{\rho\tau(1-\theta)} \Pi_{t-1}^{\rho(1-\tau)(1-\theta)} \widetilde{D}_{t+1} \right] . \tag{A9}$$

Also note that

$$\frac{\widetilde{C}_t}{\widetilde{D}_t} = \frac{C_t}{D_t} \frac{1}{P_t^{(1+\theta\omega)}} = x_t^{1+\theta\omega} , \tag{A10}$$

where $x_t \equiv X_t / P_t$. Evaluating (A8) and (A9) at the steady state leads to

$$\overline{C}_t = \frac{\frac{\theta}{\theta-1}\overline{mc}_t}{1 - \alpha \overline{q}\overline{g}^y \overline{\Pi}_t^{\theta(1+\omega)(1-\rho)}}$$

$$\overline{D}_t = \frac{1}{1 - \alpha \overline{q}\overline{g}^y \overline{\Pi}_t^{(\theta-1)(1-\rho)}}$$

Combining the two expressions above with (A3) and using (A10) leads to the steady-state restriction (8) in the main text. This restriction does not depend on τ and hence is identical to the case in Cogley and Sbordone (2008) with $\tau = 1$. Next, define $\widehat{C}_t = \ln \frac{\widetilde{C}_t}{\overline{C}_t}$, $\widehat{D}_t = \ln \frac{\widetilde{D}_t}{\overline{D}_t}$, and $\widehat{mc}_t = \ln \frac{mc_t}{\overline{mc}_t}$. Log-linearizing (A10) yields

$$(1+\theta\omega)\widehat{x}_t = (\widehat{C}_t - \widehat{D}_t) . \tag{A11}$$

Combining (A11) with (A3) and rearranging leads to an intermediate expression for $\widehat{\pi}_t$:

$$\begin{aligned} \widehat{\pi}_t &= \rho\tau \left[\widehat{\pi}_{t-1} - \widehat{g}_t^{\overline{\pi}} \right] \\ &\quad + \rho(1-\tau) \left[\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\overline{\pi}} - \widehat{g}_t^{\overline{\pi}} \right] \\ &\quad + \frac{\varphi_{0,t}}{(1+\theta\omega)} (\widehat{C}_t - \widehat{D}_t) . \end{aligned} \tag{A12}$$

We can obtain the expressions for $\widehat{C}_t$ and $\widehat{D}_t$ by log-linearizing (A8) and (A9). Combining the resulting expressions with (A11) leads to equations (1) and (2) in the main text (but with τ constrained to unity):

$$\begin{aligned}\widehat{\pi}_t &= \rho\tau(\widehat{\pi}_{t-1} - \widehat{g}_t^{\overline{\pi}}) + \rho(1-\tau)(\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\overline{\pi}} - \widehat{g}_t^{\overline{\pi}}) \\ &\quad + \lambda_t E_t(\widehat{\pi}_{t+1} - \rho\tau\widehat{\pi}_t - \rho(1-\tau)(\widehat{\pi}_{t-1} - \widehat{g}_t^{\overline{\pi}})) + \zeta_t \widehat{mc}_t + \gamma_t \widehat{D}_t + u_{\pi,t} \end{aligned} \tag{A13}$$

$$\begin{aligned}\widehat{D}_t &= \varphi_{1,t} E_t(\widehat{q}_{t,t+1} + \widehat{g}_{t+1}^{y}) \\ &\quad + \varphi_{1,t}(\theta-1)E_t\left\{\widehat{\pi}_{t+1} - \rho\tau\widehat{\pi}_t - \rho(1-\tau)(\widehat{\pi}_{t-1} - \widehat{g}_t^{\overline{\pi}})\right\} + \varphi_{1,t} E_t \widehat{D}_{t+1} ,\end{aligned} \tag{A14}$$

with the time-varying coefficients given by

$$\begin{aligned}\zeta_t &= \chi_t \varphi_{3,t} \\ \lambda_t &= \varphi_{2,t}(1 + \varphi_{0,t}) \\ \gamma_t &= \frac{\chi_t(\varphi_{2,t} - \varphi_{1,t})}{\varphi_{1,t}} \\ \chi_t &= \frac{\varphi_{0,t}}{1 + \theta\omega} \\ \varphi_{1,t} &= \alpha\overline{q}\overline{g}^{y}\overline{\Pi}_t^{(\theta-1)(1-\rho)} \\ \varphi_{2,t} &= \alpha\overline{q}\overline{g}^{y}\overline{\Pi}_t^{\theta(1+\omega)(1-\rho)} \\ \varphi_{3,t} &= 1 - \varphi_{2,t}\end{aligned}$$

Finally, iterating $\widehat{D}_t$ in (A14) forward, substituting the resulting expression for $\widehat{D}_t$ in (A13), converting real discount factors $\widehat{q}_{t+j,t+j+1}$ into nominal discount factors $\widetilde{Q}_{t+j,t+j+1}$ for ease of comparison with Cogley and Sbordone (2008), and rearranging terms yields the NKPC in DE form:

$$\begin{aligned}\widehat{\pi}_t &= \widetilde{\rho}_{1,t}^{D}\left(\widehat{\pi}_{t-1} - \widehat{g}_t^{\overline{\pi}}\right) + (1-\tau)\widetilde{\rho}_{2,t}^{D}\left(\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\overline{\pi}} - \widehat{g}_t^{\overline{\pi}}\right) \\ &\quad + \widetilde{\zeta}_t^{D}\,\widehat{mc}_t \\ &\quad + d_{1,t}^{D} E_t \widehat{\pi}_{t+1} \\ &\quad + d_{2,t}^{D} E_t \sum_{j=2}^{\infty} \varphi_{1,t}^{j-1}\widehat{\pi}_{t+j} \\ &\quad + d_{3,t}^{D} E_t \sum_{j=0}^{\infty} \varphi_{1,t}^{j}\left[\widehat{Q}_{t+j,t+j+1} + \widehat{g}_{t+j+1}^{y}\right] + \widetilde{u}_{\pi,t},\end{aligned} \tag{PC-DE}$$

where the coefficients are defined by

$$
\begin{aligned}
\widetilde{\rho}_{1,t}^{D} &= \left[\rho\tau - \lambda_t\rho(1-\tau) - \gamma_t(\theta-1)\rho(1-\tau)\varphi_{1,t}\right]/\Delta_t \\
\widetilde{\rho}_{2,t}^{D} &= \rho/\Delta_t \\
d_{1,t}^{D} &= \widetilde{d}_{1,t}^{D} + d_{3,t}^{D} \\
d_{2,t}^{D} &= \widetilde{d}_{2,t}^{D} + d_{3,t}^{D} \\
d_{3,t}^{D} &= [\gamma_t\varphi_{1,t}]/\Delta_t \\
\widetilde{\zeta}_{t}^{D} &= \zeta_t/\Delta_t \\
\Delta_t &= 1 + \rho\tau\lambda_t + \gamma_t(\theta-1)\rho\varphi_{1,t}\left\{\tau + (1-\tau)\varphi_{1,t}\right\} \\
\widetilde{d}_{1,t}^{D} &= \left[\lambda_t + \gamma_t(\theta-1)\varphi_{1,t}\left\{1 - \rho\tau\varphi_{1,t} - \rho(1-\tau)\varphi_{1,t}^2\right\}\right]/\Delta_t \\
\widetilde{d}_{2,t}^{D} &= \left[\gamma_t(\theta-1)\varphi_{1,t}\left\{1 - \rho\tau\varphi_{1,t} - \rho(1-\tau)\varphi_{1,t}^2\right\}\right]/\Delta_t
\end{aligned}
$$

Note that as in Cogley and Sbordone (2008), we use the "anticipated utility" assumption (Kreps, 1998) in deriving the NKPC in (PC-DE) so that $E_t\prod_{k=0}^{i}\varphi_{1,t+k}\widehat{h}_{t+i} = \varphi_{1,t}^{i+1}E_t\widehat{h}_{t+i}$ for any variable $\widehat{h}_{t+i}$.

Two limiting cases of (PC-DE) are worth mentioning. First, when $\tau = 1$ so that the indexation is constrained to the first lag of inflation, we have the NKPC in Cogley and Sbordone (2008), given by (3) in the main text. Second, if the prices of non-adjusting firms are fully indexed to a mixture of past inflation (first and second lags) and current trend inflation, the NKPC collapses to the case with constant coefficients and where there is no extra lead terms beyond $t+1$. Furthermore, in the constant-trends case with $\tau = 1$, one obtains the NKPC considered in Christiano, Eichenbaum, and Evans (2005).

Cross-equation restrictions Given the forecasting rule (5) and equation (PC-DE), we obtain the conditional expectation of inflation based on information at $t-2$ in the DE form as follows

$$
\begin{aligned}
\mathbf{e}_\pi'\mathbf{A}_{t-2}^2\widehat{\mathbf{z}}_{t-2} &= \widetilde{\rho}_{1,t-2}^{D}\mathbf{e}_\pi'\mathbf{A}_{t-2}\widehat{\mathbf{z}}_{t-2} + (1-\tau)\widetilde{\rho}_{2,t-2}^{D}\mathbf{e}_\pi'\widehat{\mathbf{z}}_{t-2} + \widetilde{\zeta}_{t-2}^{D}\mathbf{e}_{mc}'\mathbf{A}_{t-2}^2\widehat{\mathbf{z}}_{t-2} \\
&\quad + d_{1,t-2}^{D}\mathbf{e}_\pi'\mathbf{A}_{t-2}^3\widehat{\mathbf{z}}_{t-2} + d_{2,t-2}^{D}\varphi_{1,t-2}\mathbf{e}_\pi'\mathbf{J}_{t-2}\mathbf{A}_{t-2}^4\widehat{\mathbf{z}}_{t-2} \\
&\quad + d_{3,t-2}^{D}(\mathbf{e}_Q'\mathbf{J}_{t-2}\mathbf{A}_{t-2}^2\widehat{\mathbf{z}}_{t-2} + \mathbf{e}_{g^y}'\mathbf{J}_{t-2}\mathbf{A}_{t-2}^3\widehat{\mathbf{z}}_{t-2}), \qquad \text{(A15)}
\end{aligned}
$$

where $\mathbf{J}_t \equiv (\mathbf{I} - \varphi_{1,t}\mathbf{A}_t)^{-1}$. Hence, the vector of cross-equation restrictions is given by

$$\begin{aligned}
\mathbf{e}'_\pi \mathbf{A}^2_{t-2} &= \widetilde{\rho}^D_{1,t-2}\mathbf{e}'_\pi \mathbf{A}_{t-2} + (1-\tau)\widetilde{\rho}^D_{2,t-2}\mathbf{e}'_\pi \mathbf{I} + \widetilde{\zeta}^D_{t-2}\mathbf{e}'_{mc}\mathbf{A}^2_{t-2} \\
&\quad + d^D_{1,t-2}\mathbf{e}'_\pi \mathbf{A}^3_{t-2} + d^D_{2,t-2}\varphi_{1,t-2}\mathbf{e}'_\pi \mathbf{J}_{t-2}\mathbf{A}^4_{t-2} \\
&\quad + d^D_{3,t-2}(\mathbf{e}'_Q \mathbf{J}_{t-2}\mathbf{A}^2_{t-2} + \mathbf{e}'_{g^y}\mathbf{J}_{t-2}\mathbf{A}^3_{t-2}) \\
&\equiv \mathbf{g}^D(\boldsymbol{\mu}_{t-2}, \mathbf{A}_{t-2}, \boldsymbol{\psi}) \ . \qquad \text{(A16)}
\end{aligned}$$

Note that expectations are taken $t-2$ in order for us to remove the expectation of the trend inflation growth, which is an innovation process. For the case $\tau = 1$ considered in CS and in our main comment, it's sufficient to take expectations as of $t-1$. In this case, the resulting cross-equation restrictions are given by (6) in the main text.

Appendix B: Derivation of the NKPC in DE(∞) form

In this appendix we derive the NKPC where we impose partial model-consistent restrictions on expectations based on (A13) and (A14). The final form is equivalent to DE(∞). Once again, we use the general form with τ is unconstrained. First, define an auxiliary variable

$$\widehat{B}_t = \widehat{\pi}_t - \rho\tau(\widehat{\pi}_{t-1} - \widehat{g}^{\bar{\pi}}_t) - \rho(1-\tau)(\widehat{\pi}_{t-2} - \widehat{g}^{\bar{\pi}}_{t-1} - \widehat{g}^{\bar{\pi}}_t) \ ,$$

so that

$$E_t\widehat{B}_{t+1} = \widehat{\pi}_{t+1} - \rho\tau\widehat{\pi}_t - \rho(1-\tau)(\widehat{\pi}_{t-1} - \widehat{g}^{\bar{\pi}}_t) \ .$$

Note that the expectation above reflects the fact that $\widehat{g}^{\bar{\pi}}_t$ is an innovation process so that $E_t\widehat{g}^{\bar{\pi}}_{t+j} = 0$ for $j \geq 1$. Using this definition, we can rewrite (A13) as

$$\widehat{B}_t = \lambda_t E_t \widehat{B}_{t+1} + \zeta_t \widehat{mc}_t + \gamma_t \widehat{D}_t + u_{\pi,t}. \qquad \text{(B1)}$$

Solving forward (B1) yields

$$\widehat{B}_t = \zeta_t E_t \sum_{j=0}^{\infty} \lambda^j_t \widehat{mc}_{t+j} + \gamma_t E_t \sum_{j=0}^{\infty} \lambda^j_t \widehat{D}_{t+j} + u_{\pi,t}. \qquad \text{(B2)}$$

In deriving (B2) (and (B3) below), the "anticipated utility" assumption is used so that $E_t\zeta_{t+j}\prod_{k=0}^{j}\lambda_{t+k}\widehat{mc}_{t+j} = \zeta_t\lambda^{j+1}_t E_t\widehat{mc}_{t+j}$ and $E_t\gamma_{t+j}\prod_{k=0}^{j}\lambda_{t+k}\widehat{D}_{t+j} = \gamma_t\lambda^{j+1}_t E_t\widehat{D}_{t+j}$ for any $j > 0$. Next, solving forward

(A14), converting real discount factors into nominal ones, and rearranging lead to

$$
\begin{aligned}
\widehat{D}_t &= \varphi_{1,t} E_t \sum_{j=0}^{\infty} \varphi_{1,t}^{j} \left[\widehat{Q}_{t+j,t+j+1} + \widehat{g}_{t+j+1}^{y}\right] \\
&\quad -\kappa_{1,t}\left[\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}}\right] + \kappa_{2,t}\widehat{\pi}_t + \kappa_{3,t}\widehat{\pi}_{t+1} \\
&\quad +\kappa_{3,t} E_t \sum_{j=2}^{\infty} \varphi_{1,t}^{j-1}\widehat{\pi}_{t+j}\ ,
\end{aligned} \tag{B3}
$$

with the new coefficients defined by

$$
\begin{aligned}
\kappa_{1,t} &= (\theta-1)\rho(1-\tau)\varphi_{1,t} \\
\kappa_{2,t} &= (\theta-1)\rho\tau\varphi_{1,t} + (\theta-1)\rho(1-\tau)\varphi_{1,t}^2 \\
\kappa_{3,t} &= \theta\varphi_{1,t} - (\theta-1)\rho\tau\varphi_{1,t}^2 - (\theta-1)\rho(1-\tau)\varphi_{1,t}^3
\end{aligned}
$$

We next remove the auxiliary variables $\widehat{B}_t$ and $\widehat{D}_t$ and derive the NKPC. Using the definition of $\widehat{B}_t$, we reintroduce inflation into (B2) so that

$$
\begin{aligned}
\widehat{\pi}_t &= \rho\tau(\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}}) + \rho(1-\tau)(\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\bar{\pi}} - \widehat{g}_t^{\bar{\pi}}) \\
&\quad +\zeta_t E_t \sum_{j=0}^{\infty} \lambda_t^j \widehat{mc}_{t+j} + \gamma_t E_t \sum_{j=0}^{\infty} \lambda_t^j \widehat{D}_{t+j}\ .
\end{aligned} \tag{B4}
$$

Finally, we substitute for $\widehat{D}_{t+j}$ terms in (B4) using (B3) and rearrange the resulting expression to obtain the DE(∞) representation of NKPC:

$$
\begin{aligned}
\widehat{\pi}_t &= \widetilde{\rho}_{1,t}^{C}(\widehat{\pi}_{t-1} - \widehat{g}_t^{\bar{\pi}}) + (1-\tau)\widetilde{\rho}_{2,t}^{C}(\widehat{\pi}_{t-2} - \widehat{g}_{t-1}^{\bar{\pi}} - \widehat{g}_t^{\bar{\pi}}) \\
&\quad +\widetilde{\zeta}_t^{C} E_t \sum_{j=0}^{\infty} \lambda_t^j \widehat{mc}_{t+j} \\
&\quad +d_{0,t}^{C} E_t \sum_{k=0}^{\infty} \lambda_t^k \left[\widehat{\pi}_{t+k-1} - \widehat{g}_{t+k}^{\bar{\pi}}\right] \\
&\quad +d_{1,t}^{C} E_t \sum_{k=0}^{\infty} \lambda_t^k \widehat{\pi}_{t+k} \\
&\quad +d_{2,t}^{C} E_t \sum_{k=0}^{\infty} \lambda_t^k \widehat{\pi}_{t+k+1} \\
&\quad +d_{2,t}^{C} E_t \sum_{k=0}^{\infty} \lambda_t^k \sum_{j=2}^{\infty} \varphi_{1,t}^{j-1} \widehat{\pi}_{t+j+k} \\
&\quad +d_{3,t}^{C} E_t \sum_{k=0}^{\infty} \lambda_t^k \sum_{j=0}^{\infty} \varphi_{1,t}^{j} \left[\widehat{Q}_{t+j+k,t+j+k+1} + \widehat{g}_{t+j+k+1}^{y}\right] +\ u_{\pi,t},
\end{aligned} \tag{PC-CF}
$$

with the new coefficients defined as follows

$$\begin{aligned}
\widetilde{\rho}_{1,t}^{C} &= \rho\tau \\
\widetilde{\rho}_{2,t}^{C} &= \rho \\
\widetilde{\zeta}_{t}^{C} &= \zeta_t \\
d_{0,t}^{C} &= -\gamma_t\kappa_{1,t} \\
d_{1,t}^{C} &= -\gamma_t\kappa_{2,t} \\
d_{2,t}^{C} &= \gamma_t\kappa_{3,t} \\
d_{3,t}^{C} &= \gamma_t\varphi_{1,t}
\end{aligned}$$

Cross-equation restrictions As before, given the forecasting rule (5), the $t-2$ conditional expectation of (PC-CF) is in the form

$$\begin{aligned}
\mathbf{e}_{\pi}'\mathbf{A}_{t-2}^{2}\widehat{\mathbf{z}}_{t-2} &= \widetilde{\rho}_{1,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{A}_{t-2}\widehat{\mathbf{z}}_{t-2} + (1-\tau)\,\widetilde{\rho}_{2,t-2}^{C}\mathbf{e}_{\pi}'\widehat{\mathbf{z}}_{t-2} + \widetilde{\zeta}_{t-2}^{C}\mathbf{e}_{mc}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{2}\widehat{\mathbf{z}}_{t-2} \\
&\quad + d_{0,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}\widehat{\mathbf{z}}_{t-2} + d_{1,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{2}\widehat{\mathbf{z}}_{t-2} \\
&\quad + d_{2,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{3}\widehat{\mathbf{z}}_{t-2} + d_{2,t-2}^{C}\varphi_{1,t-2}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{4}\widehat{\mathbf{z}}_{t-2} \\
&\quad + d_{3,t-2}^{C}(\mathbf{e}_{Q}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{2}\widehat{\mathbf{z}}_{t-2} + \mathbf{e}_{g^y}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{3}\widehat{\mathbf{z}}_{t-2}), \qquad \text{(B5)}
\end{aligned}$$

where $\mathbf{K}_t \equiv (\mathbf{I} - \lambda_t\mathbf{A}_t)^{-1}$. Hence, the vector of cross-equation restrictions is given by

$$\begin{aligned}
\mathbf{e}_{\pi}'\mathbf{A}_{t-2}^{2} &= \widetilde{\rho}_{1,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{A}_{t-2} + (1-\tau)\,\widetilde{\rho}_{2,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{I} + \widetilde{\zeta}_{t-2}^{C}\mathbf{e}_{mc}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{2} \\
&\quad + d_{0,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2} + d_{1,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{2} \\
&\quad + d_{2,t-2}^{C}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{A}_{t-2}^{3} + d_{2,t-2}^{C}\varphi_{1,t-2}\mathbf{e}_{\pi}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{4} \\
&\quad + d_{3,t-2}^{C}(\mathbf{e}_{Q}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{2} + \mathbf{e}_{g^y}'\mathbf{K}_{t-2}\mathbf{J}_{t-2}\mathbf{A}_{t-2}^{3}) \\
&\equiv \mathbf{g}^{C}(\boldsymbol{\mu}_{t-2}, \mathbf{A}_{t-2}, \boldsymbol{\psi}). \qquad \text{(B6)}
\end{aligned}$$

Note that expectations are taken $t-2$ in order for us to remove the expectation of the trend inflation growth, which is an innovation process. For the case $\tau = 1$ considered in CS and in our main comment, it's sufficient to take expectations as of $t-1$. In this case, the resulting cross-equation restrictions are given by

$$
\begin{aligned}
\mathbf{e}_\pi' \mathbf{A}_{t-1} &= \widetilde{\rho}_{1,t-1}^{C} \mathbf{e}_\pi' \mathbf{I} + \widetilde{\zeta}_{t-1}^{C} \mathbf{e}_{mc}' \mathbf{K}_{t-1} \mathbf{A}_{t-1} \\
&\quad + d_{0,t-1}^{C} \mathbf{e}_\pi' \mathbf{K}_{t-2} \mathbf{I} + d_{1,t-1}^{C} \mathbf{e}_\pi' \mathbf{K}_{t-1} \mathbf{A}_{t-1} \\
&\quad + d_{2,t-1}^{C} \mathbf{e}_\pi' \mathbf{K}_{t-1} \mathbf{A}_{t-1}^{2} + d_{2,t-1}^{C} \varphi_{1,t-1} \mathbf{e}_\pi' \mathbf{K}_{t-1} \mathbf{J}_{t-1} \mathbf{A}_{t-1}^{3} \\
&\quad + d_{3,t-1}^{C} (\mathbf{e}_Q' \mathbf{K}_{t-1} \mathbf{J}_{t-1} \mathbf{A}_{t-1} + \mathbf{e}_{g^y}' \mathbf{K}_{t-1} \mathbf{J}_{t-1} \mathbf{A}_{t-1}^{2}) \\
&\equiv \mathbf{g}^{C}(\boldsymbol{\mu}_{t-1}, \mathbf{A}_{t-1}, \boldsymbol{\psi}). \qquad \text{(B7)}
\end{aligned}
$$

Appendix C: The closed form (CF), derivation and discussion

In this appendix we derive the closed-form NKPC, and the corresponding cross-equation restrictions used for estimation. In order to simplify notation, in this appendix we omit time subscripts where immaterial, and hats on variables. All expectations assume anticipated utility (Kreps 1998). We again use the more general case where the indexation is to a weighted average of two lags of inflation, with τ denotes the weight given to $t-1$ inflation.

As in the main text, we use B_t to denote non-predetermined inflation

$$
B_t \equiv \pi_t - \rho\tau\left(\pi_{t-1} - g_t^{\pi}\right) - \rho\left(1-\tau\right)\left(\pi_{t-2} - g_{t-1}^{\pi} + g_t^{\pi}\right).
$$

Then, the NKPC is given by the expression

$$
\begin{aligned}
B_t &= \lambda E_t\left\{B_{t+1}\right\} + \gamma\varphi_1\left(\theta-1\right)\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{B_{t+1+i}\right\} \\
&\quad + \zeta mc_t + \gamma\varphi_1\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{q_{t+i,t+1+i} + g_{t+1+i}^{y}\right\} + u_{\pi,t}.
\end{aligned}
$$

After rearranging the inflation terms as follows

$$
\begin{aligned}
& B_t - \lambda E_t\left\{B_{t+1}\right\} - \gamma\varphi_1\left(\theta-1\right)\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{B_{t+1+i}\right\} \\
&= \zeta mc_t + \gamma\varphi_1\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{q_{t+i,t+1+i} + g_{t+1+i}^{y}\right\} + u_{\pi,t},
\end{aligned}
$$

we add and subtract from the left-hand side of this equation the following term:

$$
\lambda\varphi_1\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{B_{t+2+i}\right\} + \varphi_1\sum_{i=0}^{\infty}\varphi_1^i E_t\left\{B_{t+1+i}\right\}.
$$

This results in the equation

$$
\begin{aligned}
& \sum_{i=0}^{\infty} \varphi_1^i E_t \left\{ B_{t+i} \right\} - \left(\lambda + \varphi_1 \left(\gamma \left(\theta - 1 \right) + 1 \right) \right) \sum_{i=0}^{\infty} \varphi_1^i E_t \left\{ B_{t+1+i} \right\} \\
& + \lambda \varphi_1 \sum_{i=0}^{\infty} \varphi_1^i E_t \left\{ B_{t+2+i} \right\} \\
= \; & \zeta m c_t + \gamma \varphi_1 \sum_{i=0}^{\infty} \varphi_1^i E_t \left\{ q_{t+i,t+1+i} + g_{t+1+i}^{y} \right\} + u_{\pi,t}.
\end{aligned}
$$

From now on we use several factorizations of the form $(1 - root \cdot F)$ in the expectational terms. We discuss later the conditions under which inverting this factor is appropriate. The first factorization extracts $(1 - \varphi_1 F)^{-1}$ from the left- and right-hand sides of the last equation to obtain

$$
\begin{aligned}
& E_t \left\{ \left(1 - \varphi_1 F \right)^{-1} \cdot \left[B_t - \left(\lambda + \varphi_1 \left(\gamma \left(\theta - 1 \right) + 1 \right) \right) B_{t+1} + \lambda \varphi_1 B_{t+2} \right] \right\} \\
= \; & \zeta m c_t + \gamma \varphi_1 E_t \left\{ \left(1 - \varphi_1 F \right)^{-1} \cdot \left[q_{t,t+1} + g_{t+1}^{y} \right] \right\} + u_{\pi,t},
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
& E_t \left\{ B_t - \left(\lambda + \varphi_1 \left(\gamma \left(\theta - 1 \right) + 1 \right) \right) B_{t+1} + \lambda \varphi_1 B_{t+2} \right\} \\
= \; & \zeta E_t \left\{ \left(1 - \varphi_1 F \right) \cdot m c_t \right\} + \gamma \varphi_1 E_t \left\{ q_{t,t+1} + g_{t+1}^{y} \right\} + u_{\pi,t}.
\end{aligned}
$$

Factoring the polynomial in the left-hand-side we obtain

$$
E_t \left\{ \left(1 - \xi_1 F \right) \left(1 - \xi_2 F \right) B_t \right\} = \zeta E_t \left\{ \left(1 - \varphi_1 F \right) \cdot m c_t \right\} + \gamma \varphi_1 E_t \left\{ q_{t,t+1} + g_{t+1}^{y} \right\} + u_{\pi,t} \tag{C1}
$$

with the roots $\{\xi_1, \xi_2\}$ given by

$$
\xi_1 + \xi_2 = \lambda + \varphi_1 + \varphi_1 \gamma \left(\theta - 1 \right), \tag{C2}
$$

$$
\xi_1 \cdot \xi_2 = \lambda \varphi_1. \tag{C3}
$$

The unique closed-form under determinacy is therefore

$$
\begin{aligned}
\pi_t = \; & \rho \tau \left(\pi_{t-1} - g_t^{\pi} \right) + \rho \left(1 - \tau \right) \left(\pi_{t-2} - g_{t-1}^{\pi} + g_t^{\pi} \right) \\
& + \zeta E_t \left\{ \sum_{j=0}^{\infty} \xi_1^j \sum_{i=0}^{\infty} \xi_2^i \left(m c_{t+i+j} - \varphi_1 m c_{t+1+i+j} \right) \right\} \\
& + \gamma \varphi_1 E_t \left\{ \sum_{j=0}^{\infty} \xi_1^j \sum_{i=0}^{\infty} \xi_2^i \left(q_{t+i+j,t+1+i+j} + g_{t+1+i+j}^{y} \right) \right\} + u_{\pi,t}.
\end{aligned} \tag{C4}
$$

This closed-form states that inflation depends on (1) indexation, (2) the expected path of real marginal costs, and (3) the real expected discounted value of future output growth. The expected path of real marginal costs enters as a quasi-difference, indicating that both the expected level and the expected rate of change of real marginal costs have an impact on inflation. While the roots $\{\xi_1, \xi_2\}$ can be complex, equation (C4) facilitates the discussion of the necessary and sufficient conditions for determinacy. At the end of this appendix we show that this closed-form is in fact a function of real-valued parameters only. For our purposes, what is important is that the cross-equation restrictions obtained from (C4), to which we now turn, involve only real-valued coefficients.

To derive the cross equation restrictions, we use the forecasting rule (5) to form expectations. Taking expectations of (C4) conditional on information available at time $t-2$, we obtain

$$\begin{aligned} e'_\pi \mathbf{A}^2_{t-2} &= \rho\tau e'_\pi \mathbf{A}_{t-2} + \rho (1-\tau) e'_\pi \mathbf{I} \\ &+ \zeta_{t-2} e'_{mc} \mathbf{A}^2_{t-2} \left(\mathbf{I} - \varphi_{1,t-2}\mathbf{A}_{t-2}\right) \mathbf{W}^{-1}_{t-2} \\ &+ \gamma_{t-2}\varphi_{1,t-2} \left[\widetilde{e}'_{q,t-2}\mathbf{I} + e'_y \mathbf{A}_{t-2}\right] \mathbf{A}^2_{t-2}\mathbf{W}^{-1}_{t-2}, \end{aligned}$$

where the invertible matrix $\mathbf{W}_t$ is defined as

$$\begin{aligned} \mathbf{W}_t &\equiv \left(\mathbf{I} - \xi_{1,t}\mathbf{A}_t\right)\left(\mathbf{I} - \xi_{2,t}\mathbf{A}_t\right) \\ &= \left[\left(\mathbf{I} - \lambda_t \mathbf{A}_t\right)\left(\mathbf{I} - \varphi_t \mathbf{A}_t\right) - \gamma_t \varphi_{1,t} (\theta - 1) \mathbf{A}_t\right], \end{aligned}$$

with the last equality following immediately from (C1)-(C3), and the forecasting rule, and the vector $\widetilde{e}'_{q,t}$ is given by

$$\widetilde{e}'_{q,t} = e'_Q + e'_\pi \mathbf{A}_t.$$

The vectors e'_x are row vectors with zeros everywhere except for a single one in the position of variable x_t in $\mathbf{z}_t$. The vector $\widetilde{e}'_{q,t-2}$ selects the expected real discount factor by adding expected inflation to the nominal discount factor, which is the discount factor used in the estimated VAR. The closed-form for inflation cannot depend on expected future nominal variables, as these would contain expected future inflation. For this reason, it is necessary to add expected inflation back to the nominal discount factor, in order to obtain the expected real discount factor.

From expression (C4) it is possible to extract the necessary and sufficient conditions for the validity of the closed-form estimates (see also Appendix D). These conditions are

$$\begin{aligned} \|\varphi_1 \mathbf{A}\| &< 1 \\ \|\xi_i \mathbf{A}\| &< 1 \text{ for } i = 1, 2. \end{aligned}$$

Based on various estimates in the main text, these conditions appear to be satisfied most of the time in the data. A property of $\{\xi_1, \xi_2\}$ worth noting is that as long as $\gamma_t \varphi_{1,t} (\theta - 1) \geq 0$, which occurs in over 99 percent of the estimates, we have that

$$\begin{aligned} \xi_{1,t} &\geq \lambda_t \\ \xi_{2,t} &\leq \varphi_{1,t}. \end{aligned}$$

This can be shown from simple but tedious algebra based on the expression of these roots

$$\{\xi_1, \xi_2\} = \frac{2\lambda\varphi_1}{\lambda + \varphi_1 + \varphi_1 \gamma (\theta - 1) \pm \sqrt{(\lambda + \varphi_1 + \varphi_1 \gamma (\theta - 1))^2 - 4\lambda\varphi_1}}.$$

We also note that when $\gamma_t \varphi_{1,t} (\theta - 1) = 0$ we obtain

$$\mathbf{W}_t^{-1} = (\mathbf{I} - \varphi_t \mathbf{A}_t)^{-1} (\mathbf{I} - \lambda_t \mathbf{A}_t)^{-1},$$

which shows that in this case $\xi_1 = \lambda_t$, and $\xi_2 = \varphi_{1,t}$. Hence, conditional on $\left\|\varphi_{1,t} \mathbf{A}_t\right\| < 1$, the necessary and sufficient condition would be given by

$$\|\lambda_t \mathbf{A}_t\| < 1,$$

which is the same condition discussed in the main text for the existence of CF estimators (given that $\left\|\varphi_{1,t} \mathbf{A}_t\right\| < 1$). In our closed-form (unrestricted) estimation we obtain that $\left|\gamma_t \varphi_{1,t} (\theta - 1)\right|$ has a median of $8.7x10^{-4}$, a mode of $2.8x10^{-6}$, and is below $3x10^{-3}$ for all ensembles.[1] In short, the necessary and sufficient conditions for the existence of the closed-form are extremely close to the conditions affecting the CF specifications. Also, as discussed in the main text, the conditions for the existence of our CF estimators are always met whenever the closed form solution exists.

To conclude, we show that even when the roots $\{\xi_1,\ \xi_2\}$ are complex, the closed-form in (C4) only depends on real-valued coefficients.[2] To see this, note that

$$\sum_{j=0}^{\infty} \xi_1^j \sum_{i=0}^{\infty} \xi_2^i E_t \{mc_{t+i+j}\} = \sum_{j=0}^{\infty} \delta (\xi_2 + \xi_1, \xi_2 \xi_1, j) E_t \{mc_{t+j}\}$$

[1] Negative values of $\gamma_t \varphi_{1,t} (\theta - 1)$ are typically due to $\gamma_t < 0$, which occurs for 1.17 percent of the 765,948 point estimates. We have only one negative estimate of $\varphi_{1,t}$, and the estimation is restricted to produce estimates of $\theta > 1$.

[2] This occurs for less than 0.06 percent of our estimates.

where δ is a real-valued function of real arguments.[3] This function is given by

$$\delta\left(\xi_2+\xi_1,\ \xi_2\xi_1,\ j\right) = \begin{cases} \left(\xi_2+\xi_1\right)^j \quad \text{for } j=0,\ 1. \\ \left(\xi_2+\xi_1\right)^j + \sum_{i=1}^{M(j)} (-1)^j \binom{j-i}{i} \left(\xi_2+\xi_1\right)_1^{j-2i} \left(\xi_2\xi_1\right)^i \quad \text{for } j \geq 2 \end{cases}$$

$$M(j) = \begin{cases} \frac{j}{2} & \text{for } j \text{ even,} \\ \frac{j-1}{2} & \text{for } j \geq 3 \text{ and } j \text{ odd,} \end{cases}$$

where $\binom{j-i}{i}$ is the binomial coefficient ${}_{j-i}C_i$, and $M(j)$ determines the number of elements in the right-hand side summation. Despite its complicated appearance, this function essentially provides the moving average coefficients from inverting a stationary $AR(2)$ process of the form y_t $(1-\xi_2 L)(1-\xi_1 L) = e_t$.

Relationship between cross-equation restrictions of the DE specification and the closed-form (CF) solution Next, we discuss the relationship between the cross-equation restrictions of the DE specification and the closed-form solution. The cross-equation restriction errors for the DE specification with time-varying trends can be written as

$$\begin{aligned} \mathbf{F}^D\left(\mathbf{A}_t,\psi,t\right) \equiv\ & e_\pi' \mathbf{B}_t^* \left(\mathbf{I}-\lambda_t \mathbf{A}_t\right) \\ & -\zeta e_{mc}' \mathbf{A}_t \left(\mathbf{I}-\varphi\mathbf{A}\right) \\ & -\gamma_t \varphi_{1,t} \left(\theta-1\right) e_\pi' \mathbf{B}_t^* \mathbf{A}_t \\ & -\gamma_t \varphi_{1,t} \left[\widetilde{e}_{q,t}' \mathbf{I} + e_y' \mathbf{A}\right], \end{aligned} \tag{C5}$$

$$\text{with } \mathbf{B}_t^* \equiv \left(\mathbf{A}_t^2 - \rho\tau\mathbf{A}_t - \rho\left(1-\tau\right)\mathbf{I}\right)\left(\mathbf{I}-\varphi\mathbf{A}\right),$$

although in the main text we provide the expression closest to Cogley and Sbordone for ease of comparison (this also motivates the particular presentation of the CF cross-equation restrictions in the main text). The differences are only due to the grouping of specific terms in the equation, and for this appendix we adopt the expression (C5) to facilitate our discussion here. This expression can be directly derived from the difference-equation that results from quasi-differencing forward the Phillips curve defined by (A13) and (A14), with factor φ_1.

The cross-equation restriction errors of the ECF specification are then given by

$$\mathbf{F}^{ECF}\left(\mathbf{A}_t,\boldsymbol{\psi},t\right) = \mathbf{F}^D\left(\mathbf{A}_t,\boldsymbol{\psi},t\right)\left(\mathbf{I}+\xi_1\mathbf{A}+\left(\xi_1\mathbf{A}\right)^2+...\right)\left(\mathbf{I}+\xi_2\mathbf{A}+\left(\xi_2\mathbf{A}\right)^2+...\right)$$

[3] The arguments $(\xi_1+\xi_2)$ and $(\xi_1\xi_2)$ are given in (C2) and (C3) above as a real-valued function of the parameter values, which are always real. The third argument is a non-negative integer.

which illustrates that $\mathbf{F}^{ECF}(\mathbf{A}_t, \boldsymbol{\psi}, t)$ is a reweighting of the difference equation errors, as in the constant coefficients case. And that they can be obtained as a weighted sum of the restrictions imposed by the model on expectations of future variables. In this sense, the only difference with the constant coefficients case is this involves a double summation. Nonetheless, since the idea behind imposing model-consistency restrictions on expectations in our application is to explicitly build the link between expected period $t+j$ inflation and expected period $t+j$ marginal costs, imposing this link for m periods is equivalent to

$$\mathbf{F}^{m}(\mathbf{A}_t, \boldsymbol{\psi}, t) = \mathbf{F}^{D}(\mathbf{A}_t, \boldsymbol{\psi}, t)\left(\mathbf{I} + \delta_{t,\xi}(1)\,\mathbf{A}_t + ... + \delta_{t,\xi}(m)\,\mathbf{A}_t^{m}\right)$$

with $\delta_{t,\xi}(j) \equiv \delta\left(\xi_{1,t} + \xi_{2,t},\ \xi_{1,t}\xi_{2,t},\ j\right)$, and hence $\mathbf{F}^{m} \rightarrow \mathbf{F}^{ECF}$ when $m \rightarrow \infty$.

Appendix D: Misspecification of the first-stage VAR in the presence of sunspots

In this appendix we show that the VAR representation (4) in the main text with a finite number of lags and i.i.d. shocks implies that there is a unique forward solution to the model. In particular, we show that indeterminacy of the model solution yields to a reduced-form representation that involves a moving average error term. In such a setting, a reduced-form VAR representation with i.i.d. shocks entails an infinite number of lags.[4] Therefore, estimation of the finite-order VAR in (4) results in biased and inconsistent coefficient estimates (truncation bias). This discussion is relegated to an appendix because its content draws on a vast literature on solutions to rational expectations models; see for instance the textbook treatment in Pesaran (1989) and the references therein.

For ease of exposition and without loss of generality, we consider a simple setup with only two structural equations.[5] The first equation is a purely forward-looking version of the NKPC with zero (constant) trend inflation:

$$\pi_t = \Lambda E_t \pi_{t+1} + \zeta mc_t + u_{\pi,t}, \tag{D1}$$

where $\zeta \geq 0$ is the elasticity of inflation with respect to marginal costs, $\Lambda > 0$ determines the extent of forward-looking behavior, and $u_{\pi,t}$ denotes the structural i.i.d. shock associated with the NKPC.

[4] For a discussion of this non-invertibility problem see Fernandez-Villaverde et al. (2007).

[5] The arguments here carry over with some minor adjustments to the anticipated utility framework with with time-varying trends (or parameters). Also, note that our discussion focuses on the issue of determinacy and assumes for convenience that a stationary solution is feasible.

The presence of this structural shock is critical for modern DSGE models and is at the heart of the discussion in this appendix.[6] The second equation closes the model, and states that marginal costs follow a simple univariate autoregressive process

$$mc_t = \chi mc_{t-1} + u_{mc,t}, \tag{D2}$$

where $1 \geq \chi > 0$ and $u_{mc,t}$ is also a structural i.i.d. shock.

Before analyzing the conditions for determinacy, notice that regardless of the value of $\Lambda > 0$, it is always possible to express (D1) backwards as

$$\pi_t = \frac{1}{\Lambda}\pi_{t-1} - \frac{\zeta}{\Lambda}mc_{t-1} - \frac{1}{\Lambda}u_{\pi,t-1} + \left[\pi_t - E_{t-1}\pi_t\right]. \tag{D3}$$

In a determinate solution the model equations impose just enough conditions to uniquely determine the forecast error $[\pi_t - E_{t-1}\pi_t]$ as a linear combination of new information revealed at time t by the two structural shocks $u_{\pi,t}$, and $u_{mc,t}$. When there are not enough conditions to uniquely pin down the forecast error $[\pi_t - E_{t-1}\pi_t]$, rational expectations only dictate that the error be unpredictable as of time $t-1$, and as a result, this error could be partly driven by a sunspot.[7] Whether the stationary solution for the system in (D1) and (D2) is indeterminate cannot be established just by looking at the parameters of the NKPC. In fact, in this stylized framework whether a stationary solution involves sunspots or not is uniquely determined by the parameter interaction $\Lambda\chi$.

Consider first the case in which $\Lambda\chi < 1$. Then a unique stationary forward-looking solution for the NKPC exists and takes the following form:

$$\begin{aligned}
\pi_t &= \zeta \sum_{i=0}^{\infty} \Lambda^i E_t mc_{t+i} + u_{\pi,t} \\
\Leftrightarrow \pi_t &= \zeta mc_t \sum_{i=0}^{\infty} (\Lambda\chi)^i + u_{\pi,t} \\
\Leftrightarrow \pi_t &= \frac{\zeta}{1-\Lambda\chi} mc_t + u_{\pi,t}.
\end{aligned} \tag{D4}$$

In constrained reduced-form, the last expression can therefore be rewritten as

$$\pi_t = \frac{\zeta\chi}{1-\Lambda\chi}mc_{t-1} + u_{\pi,t} + \frac{\zeta}{1-\Lambda\chi}u_{mc,t}. \tag{D5}$$

Equations (D2) and (D5) form a constrained reduced-form VAR with serially uncorrelated shocks. In order to estimate the deep parameters of the model it is possible to minimize the

[6] This shock can be interpreted as capturing potential misspecifications in the relationship or a markup shock.

[7] See Pesaran (1989, Chapter 5) for a discussion of the different methods to refine the solution set under indeterminacy.

distance between the constrained VAR given by (D2) and (D5) and an estimated unconstrained reduced-form VAR as in (4). In the specific case of our application, we try to match the cross-equation restrictions implied by model expectations, with forecasts obtained from the estimated unrestricted VAR. For this exercise to be correct, the estimated unrestricted VAR in (4) and the restricted VAR implied by (D2) and (D5) need to share two crucial characteristics: (1) a finite number of lags, and (2) i.i.d. shocks.

Using (D4) it is possible to eliminate $u_{\pi,t-1}$ in (D3), to obtain

$$\pi_t = \frac{\zeta\chi}{1-\Lambda\chi} mc_{t-1} + (\pi_t - E_{t-1}\pi_t). \tag{D6}$$

Next, equations (D5) and (D6) can be used to solve for the forecast error as a linear combination of the structural shocks:

$$\pi_t - E_{t-1}\pi_t = u_{\pi,t} + \frac{\zeta}{1-\Lambda\chi} u_{mc,t}.$$

Hence, in the determinate solution the time t values of inflation, marginal costs, and expected inflation, are all uniquely determined by the two structural shocks in the model. At most two of these variables can be linearly independent. This property is what allows to properly form inflation expectations using a VAR where the only right-hand-side variables are inflation and marginal costs. As we discuss next, this property does not hold in the presence of sunspots.

The indeterminate (sunspot) solution arises when $\Lambda\chi \geq 1$. In this case, the stationary forward-looking solution we just discussed is no longer feasible because equations (D4), (D5) and (D6) do not hold. This implies that the forecast errors cannot be uniquely determined by the new time t information provided by $u_{\pi,t}$ and $u_{mc,t}$. Following Lubik and Schorfheide (2003), the forecast error in this case can be written as

$$\pi_t - E_{t-1}\pi_t = M_1 u_{\pi,t} + M_2 u_{mc,t} + h_t. \tag{D7}$$

where h_t is an arbitrary martingale difference (the sunspot shock), and M_1 and M_2 are parameters that are not pinned down by the structural parameters of the model. As a result, the parameters M_1 and M_2 cannot be guaranteed to take any specific value, although in principle they can be estimated (see Lubik and Schorfheide 2004).[8]

Replacing the forecast error in (D3) by (D7) yields the following constrained reduced-form

[8] For estimation purposes, they cannot be assumed to be zero, or to equal their determinacy values $M_1 = 1$, and $M_2 = \zeta/(1-\Lambda\chi)$.

representation of inflation

$$\pi_t = \frac{1}{\Lambda}\pi_{t-1} - \frac{\zeta}{\Lambda}mc_{t-1} + M_2 u_{mc,t} + \xi_t \qquad \text{(D8)}$$

$$\xi_t = \left[M_1 u_{\pi,t} - \frac{1}{\Lambda}u_{\pi,t-1}\right] + h_t. \qquad \text{(D9)}$$

Under determinacy, equation (D4) implies that the error $u_{\pi,t-1}$ in (D9) can be replaced by a linear combination of π_{t-1} and mc_{t-1}, and as a result inflation does not contain a moving average error term. But in the indeterminate case equation (D4) does not hold, and the moving average error in (D9) becomes an important problem for estimation. More specifically, the equilibrium under indeterminacy is spanned by three independent shocks ($u_{\pi,t}$, $u_{mc,t}$ and ξ_t), and as a result marginal costs, inflation, and expected inflation are all linearly independent.[9] To avoid misspecification, a finite-order VAR would need to include all three variables: inflation, marginal costs, and most importantly inflation expectations themselves.[10]

It follows from this discussion that the dynamics of marginal costs and inflation given by (D2) and (D8) cannot have an unrestricted VAR representation such as (4), because (4) does not allow for a moving average error component. In other words, the problem with the unrestricted VAR in (4) in the main text is the omission of a variable that belongs in the model. In the presence of sunspots the estimated VAR parameters are biased and inconsistent, since the lagged endogenous variables at $t-1$ are correlated with the moving average error term. Moreover, even if consistent estimates for the VAR in (4) part of the process were available, it would still be incorrect to use (4) to proxy for one-period-ahead expectations.[11] The error term $u_{\pi,t-1}$ would contribute valuable information to form time $t-1$ inflation expectations beyond what would be contributed by inflation and marginal costs.

Overall, this simple example shows that assuming the VAR representation (4) in the main text as an unconstrained reduced-form for the NKPC model amounts to imposing the existence of a stable forward-looking solution, i.e. $\Lambda\chi < 1$. If $\Lambda\chi \geq 1$, the reduced-form VAR representation contains a moving average error term. Alternatively, the VAR can be written as a function of serially uncorrelated shocks only by introducing an infinite number of lags. Otherwise estimating (4) with a finite number of lags results in truncation bias in the estimated coefficients. This

[9] Note that as long as (D4) does not hold, inflation is affected by the moving average ξ_t even if we assume $h_t = 0$ for all t. More generally, the problems posed by ξ_t for the invertibility of the DSGE model are present even if $M_1 = 0$.

[10] This is, in essence, the non-invertibility problem that arises whenever the VAR has to incluede variables that are unobservable to the econometrician (see Fernandez-Villaverde et al. 2007).

[11] In order to obtain these consistent estimates the empirical model could be a $VARMA(p, q)$ model. In models that involve additional leads and lags of inflation, the problem is how to determine the correct values of p and q.

bias would undermine the validity of the second-stage, whereby estimates of the parameters of the NKPC are obtained by imposing model-consistent cross-equation restrictions on the estimated VAR. Moreover, because the estimation of (4) in the first-stage imposes the assumption that no sunspots are present in the equilibrium, it is important that the second-stage estimates are supportive of this assumption. Evidence to the contrary would invalidate the estimation because of an incongruence between the two estimation stages. In our application, determinacy of the system hinges on the largest eigenvalues of $\widehat{\lambda}_t\widehat{\mathbf{A}}_t$. Figure 1 in our comment illustrates that this is largely not a problem for our CF estimates, but could be an important problem for the DE estimates in Cogley and Sbordone.

Two additional issues are worth noting here. First, truncation bias would imply that the estimated errors from the VAR follows an infinite-order moving average (see Fernandez-Villaverde et al. 2007). Second, even if a finite-order VAR might provide a relatively good approximation to the dynamics of the true infinite order VAR for specific shocks, this is ultimately irrelevant for our estimation because we exploit cross-equation restrictions.[12] In particular, the dynamics of the variables are determined by linear combinations of the columns of the VAR, while the cross equation restrictions select very specific elements of the estimated VAR matrix to determine the structural parameters. Even if the linear combinations of the inconsistently estimated VAR matrix may provide relatively good impulse responses for specific shocks, the cross-equation restrictions hinge much more critically on each point estimate involved in a cross-equation restriction. In this sense, for the purpose of minimum-distance estimation as carried out in our paper and in Cogley and Sbordone (2008), the warning in Fernandez-Villaverde et al. (2007) against simply assuming that a finite-order VAR appropriately replicates model dynamics is particularly relevant.

[12] See Sims and Zha (2006) and Sims (2009) for applications where the non-invertibility of a DSGE model does not pose a particularly serious issue for estimation of a finite-order VAR.