

Reicher, Christopher Phillip

Conference Paper

Evaluating misspecification in DSGE models using tests for overidentifying restrictions

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Moment Matching: Application of an Alternative Estimation Approach in Macroeconomics, No. C20-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Reicher, Christopher Phillip (2013) : Evaluating misspecification in DSGE models using tests for overidentifying restrictions, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Moment Matching: Application of an Alternative Estimation Approach in Macroeconomics, No. C20-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/79955>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Evaluating misspecification in DSGE models using tests for overidentifying restrictions

Christopher Phillip Reicher*

Institut für Weltwirtschaft, Kiel, Germany

This version: February 1, 2013

Abstract

In this paper I discuss the estimation of the process governing the structural shocks (or wedges) to a DSGE model, arguing that a well-specified model would satisfy certain sets of moment conditions. Based on tests for overidentifying restrictions, I compare three specifications of the Taylor rule within a simple New Keynesian model. I find that a rule which allows for the Fed to respond to four lags of inflation shows less evidence of misspecification than one where the Fed responds only to contemporaneous inflation. Raising the coefficient on the output gap to 1 instead of 0.5 gives more ambiguous results.

*Phone: +49 (0)431 8814 300. Email: christopher.reicher@ifw-kiel.de. Address: Institut für Weltwirtschaft, Hindenburgufer 66, 24105 Kiel, Germany. I wish to thank Steffen Ahrens, Martin Plödt, Jens Boysen-Hogrefe, Vincenzo Caponi, Dominik Groll, and Henning Weber for their patient advice and valuable feedback. All errors are mine.

JEL: C12, C22, E52

Keywords: Overidentifying restrictions, wedges, vector autoregression, misspecification

1 Introduction

In this paper I discuss the evaluation of the evidence in favor of misspecification for a DSGE model using tests for overidentifying restrictions, where the overidentifying restrictions represent a hypothesis that the model is not misspecified along some set of dimensions. First I show how an unrestricted VAR process governing the shocks (or wedges) in a DSGE model can be estimated by running a VAR on observables and then backing out the implied law of motion for the wedges given a set of deep structural parameters for the model. Then I propose a set of absolute and relative misspecification test statistics based on χ^2 tests for overidentifying restrictions, based on the heuristic notion that a well-specified model satisfies certain zero restrictions. In practical terms, some combinations of parameters which govern the law of motion for the wedges—particularly wedge processes which feature strong relationships among different "structural" wedges—are more indicative of misspecification than other combinations of parameters. These test statistics can provide practical guidance toward building better models.

I use the specification of the Taylor rule in a three-equation New Keynesian model as a concrete example to see how tests for overidentifying restrictions could be used to improve model specification. In particular, I test to see to what extent the data reject the null hypothesis that the errors for a set of three model equations are not correlated across equations, for a number of time horizons. Based on the model misspecification statistics, I show that a Taylor rule where monetary authorities respond to one year's worth of cumulative inflation requires assumptions about the underlying shock process which are less unbelievable along the specified dimensions than a Taylor rule where monetary authorities respond only to current quarterly inflation. I also compare a model with a Taylor rule coefficient on output of 1 with a rule with a coefficient of 0.5, and I find that a larger coefficient seems to be better-specified when looking at contemporaneous data but worse-specified when looking at lagged data. In summary, it is possible to test for certain types of model misspecification in DSGE models given that the data follow a VAR process and to make improvements to

model specification based on the results from these tests.

There is already a small literature which seeks to estimate the wedges in dynamic economic models using generalized shock processes, given a fixed set of deep model parameters. Relative to that literature, I present a computationally simpler way to estimate the process governing the evolution of model errors (or wedges), which can then be applied to the detection and correction of model misspecification. This contribution is potentially useful in its own right. Chari, Kehoe, and McGrattan (2007) estimate a set of wedges in an RBC framework, and Šustek (2011) estimates a version of the Smets-Wouters (2007) model with a generalized driving process governing the model wedges. Šustek, for instance, has to resort to simulated annealing in order to approximate the wedge process. Cúrdia and Reis (2011) attempt to estimate both the deep parameters and the parameters governing the errors to a large DSGE model with VAR(1) errors through the use of computationally intensive Bayesian methods. It turns out that there is no need to use computationally intensive methods to estimate the underlying driving process if the data and driving processes are both VARs, since it is possible to estimate the VAR which governs the data and then back out the VAR which governs the wedges. This result makes it possible to derive an estimated law of motion for the wedges in a fast and accurate way.

While in this paper I adopt a classical approach, this paper is also related to the Bayesian model evaluation literature. An and Schorfheide (2007) extensively discuss the use of Bayesian methods such as posterior predictive checks and odds ratios to compare potentially misspecified models; the model misspecification statistics discussed in this paper could be seen as applying the logic of posterior predictive checks to a more classical setting. Additionally, del Negro and Schorfheide (2004) discuss how the use of a DSGE model as a prior distribution for a VAR model could improve forecasting performance and policy predictions, while del Negro and Schorfheide (2009) discuss further ways in which misspecification can be taken into account when doing policy analysis with DSGE models. In contrast with the Bayesian literature, I adopt a classical approach for the sake of computational simplicity,

though this paper retains the spirit of much of the Bayesian literature. Mechanically, the test statistics used in this paper are the same test statistics as those used in the generalized method of moments (GMM) literature as well as in the structural equation modeling (SEM) literature. These test statistics do not represent a coherent approach to the estimation of deep parameters in the way that maximum likelihood, GMM, and Bayesian estimation do; rather, I seek to develop a set of relatively simple tools which can be used to help diagnose misspecification and to make it possible to build better models. Future work could see to what extent evidence penalties for misspecification could be integrated into a more general estimation strategy.

The layout of this paper is as follows. First, I show how to back out an estimated law of motion for model wedges from an estimated law of motion for the data, for a given set of deep DSGE model parameters. Then I show how to set up a set of overidentifying restrictions tests on undesirable parameter combinations, and I show how these overidentifying restrictions tests could be used to improve the specification of the Taylor rule in a three-equation New Keynesian model. I conclude with a discussion regarding further possible directions for research.

2 The mapping between data and wedges

2.1 The mapping from wedges to observables

In this section I show that an unrestricted VAR process for the shocks to an economic model results in a VAR process for the observables. The following section shows the reverse. This section and the next section closely follow the setup used by Reicher (2013) to discuss identification issues in DSGE models with generalized shock processes, although this paper goes on to show how to computationally derive the estimated law of motion for the wedges and to perform specification tests.

A dynamic economic model is governed by a parameter set θ . Elements of θ may include

parameters which govern production technology, labor supply, frictions in price and wage formation, adjustment costs, and so on. A set of mean-zero exogenous structural wedges w_t of rank k follows a VAR process with the law of motion:

$$w_t = \sum_{i=1}^{\infty} F_i w_{t-i} + \zeta_t, \text{ where } E\zeta_t \zeta_t' = \Sigma_w. \quad (1)$$

A complete linearized, de-meaned DSGE model including wedges and observables can be represented using the notation of Sims (2002), treating the law of motion (1) of the wedges as given:

$$\Gamma_{0,0}x_{t+1} = \Gamma_{1,0}x_t + \Pi_0\eta_{t+1} + \Psi_0\zeta_{t+1}, \quad (2)$$

where the matrices $\Gamma_{0,0}$, $\Gamma_{1,0}$, Π_0 , and Ψ_0 are functions of deep model parameters and of the law of motion of the wedges given by (1). The matrix x_t contains the observables z_t , the wedges w_t , and any other auxiliary variables included in the model. Endogenous expectational errors, which are functions of the shocks to the wedges, are given by η_{t+1} . I assume that ζ_t , w_t , and z_t have the same rank k , and that the model is stationary and locally determinate. If the rank of w_t exceeds the rank of z_t , then the mapping from z_t into w_t is not unique. If the rank of z_t exceeds the rank of w_t , then the model is stochastically singular.

The observables z_t are linked to the system through the observation equation $z_t = Hx_t$. The law of motion for the wedges (1) and the model equations form the rows of (2). The model has a solution of the form:

$$x_t = A_0x_{t-1} + B_0\zeta_t, \quad (3)$$

which usually has to be solved for numerically. I assume that the resulting law of motion is

strictly stationary and determinate. Iterating (3) forward yields:

$$z_t = Hx_t = H \sum_{i=0}^{\infty} A_0^i L^i B_0 \zeta_t. \quad (4)$$

The expectational error for z_t , denoted by ε_t , equals $HB_0\zeta_t$, assuming that $H\Lambda B_0$ must be of full rank for all square matrices Λ themselves of full rank. This is equivalent to assuming that the model and data are relevant for each other. Substituting these relationships into (4) gives the potentially infinite-order MA and VAR processes which govern the evolution of the data, respectively:

$$z_t = H \sum_{i=0}^{\infty} A_0^i L^i B_0 (HB_0)^{-1} \varepsilon_t, \quad (5)$$

and

$$z_t = \sum_{i=1}^{\infty} \Phi_i z_{t-i} + \varepsilon_t, \text{ where } E\varepsilon_t \varepsilon_t' = \Sigma_z = HB_0 \Sigma_w B_0' H', \quad (6)$$

with some set of VAR coefficients given by Φ and a covariance matrix for the errors given by Σ_z . This result is an exact version of the result utilized by del Negro and Schorfheide (2004, 2009) to form the basis of the DSGE-VAR literature. The VAR in observables be of infinite order as shown by Ravenna (2007), particularly when there are endogenous state variables. This result implies that given a well-specified linearized DSGE model, it makes sense in the first place to look at situations where the observables themselves follow a VAR process.

2.2 The mapping from observables to wedges

This section dervies the mapping between a VAR in observables and a VAR in the wedges, using the same logic from the previous section. To show that a given process governing (6) implies a process governing (1), I again represent the model using the notation of Sims (2002)

in equation (2), but this time I treat the law of motion (6) as given instead of (1). The steps otherwise proceed as before. The variables x_t again contain the observables z_t , the wedges w_t , and any other auxiliary variables included in the model. I again assume that ε_t , w_t , and z_t have the same rank k and that the model is stationary and locally determinate. The system in the form of Sims (2002) now takes the form:

$$\Gamma_{0,1}x_{t+1} = \Gamma_{1,1}x_t + \Pi_1\eta_{t+1} + \Psi_1\varepsilon_{t+1}, \quad (7)$$

where the matrices $\Gamma_{0,1}$, $\Gamma_{1,1}$, Π_1 , and Ψ_1 are functions of deep model parameters and of the law of motion of the *observables* given by (6). The wedges w_t are linked to the system through the observation equation $w_t = Dx_t$. The augmented model has a solution of the form:

$$x_t = A_1x_{t-1} + B_1\varepsilon_t, \quad (8)$$

which usually has to be derived numerically. Iterating (8) forward yields:

$$w_t = Dx_t = D \sum_{i=0}^{\infty} A_1^i L^i B_1 \varepsilon_t. \quad (9)$$

The expectational error for w_t , denoted by ζ_t , equals $DB_1\varepsilon_t$, assuming that $D\Lambda B_1$ must be of full rank for all square matrices Λ themselves of full rank. This is a relevance assumption similar to the relevance assumption made in the previous section. Substitution gives the infinite-order MA process which governs the evolution of the wedges:

$$w_t = D \sum_{i=0}^{\infty} A_1^i L^i B_1 (DB_1)^{-1} \zeta_t, \quad (10)$$

which, in lag notation, can be written as following $w_t = D(I - A_1 L)^{-1} B_1 (D B_1)^{-1} \zeta_t$, from which the VAR coefficients in the form of (1) may in principle be derived analytically.

In practice, lag notation is difficult to work with in a computational environment. Instead, the system governed by (1) can be approximated by first estimating (6) as a P -lag VAR for a large enough P (typically 4 for quarterly data), deriving equations (7) and (8) for a given θ , and then calculating the implied contemporaneous covariance matrix of w_t , denoted by $\Sigma_{w:0,0}$, the implied covariance matrix of the first P lags of w_t , denoted by $\Sigma_{w:1,P}$, and then also calculating the implied partial covariance matrix between w_t and its first P lags, denoted by $\Gamma_{w:1,P}$. These objects are of dimension k by k , kP by kP and kP by k , respectively. Letting F equal the k by kP vector of the horizontally stacked individual matrices F_i from (1), the matrices F and Σ_w can be approximated using the population regression formula $F' = (\Sigma_{w:1,P})^{-1} \Gamma_{w:1,P}$ and the covariance formula $\Sigma_w = \Sigma_{w:0,0} - F \Sigma_{w:1,P} F'$, respectively. Alternately, a system given by (8) may be analyzed directly with respect to its covariance structure.

The relatively straightforward mapping from the data generating process to the wedge generating process makes it possible to examine the process which governs the wedges. It is far easier to obtain this mapping than to nonlinearly optimize the likelihood function using the mapping from F_i and Σ_w to the observables, conditioning on θ . Knowing the mapping from the data generating process to the wedge generating process makes it possible to discuss model misspecification in a heuristic way. If the process which governs the wedges appears to violate a set of zero restrictions, then the econometric evidence can be said to reflect the possibility of misspecification along those dimensions. It is then possible to test for certain types of misspecification by performing tests for overidentifying restrictions on the model implied by (1). The next section gives a concrete example of how this could be done in the context of a simple New Keynesian DSGE model.

3 A simple DSGE model

Here, I illustrate how the results from the previous section could be used to detect misspecification in the Taylor rule in a simple 3-equation New Keynesian model. The output gap y_t is related to the inflation gap π_t through an aggregate supply equation where the parameter κ reflects the effect of inflation on output, and β is the rate of time preference. Current output is related to future output, future inflation, and current interest rates through an aggregate demand equation, where the parameter σ governs the willingness of consumers to substitute across time. Interest rates are governed by a Taylor rule which relates interest rates to inflation, output, and the lagged interest rate through the Taylor rule coefficients ϕ_π , ϕ_y , and ρ_i , respectively. The system, with wedges w_t representing equation errors, is expressed by the following three equations:

$$y_t = \kappa\pi_t - \kappa\beta E_t\pi_{t+1} + w_t^s; \quad (11)$$

$$y_t = -\frac{1}{\sigma}(i_t - E_t\pi_{t+1}) + E_ty_{t+1} + w_t^d; \quad (12)$$

and

$$i_t = (1 - \rho_i)\phi_\pi\pi_t + (1 - \rho_i)\phi_y y_t + \rho_i i_{t-1} + w_t^i. \quad (13)$$

The wedges w_t^s , w_t^d , and w_t^i represent, respectively, reduced-form disturbances to aggregate supply, aggregate demand, and monetary policy.

The system written in the canonical form (7) takes the following form when the observables follow a VAR(1) as in (6):

$$\begin{aligned}
& \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \kappa\beta & 0 & 0 & 0 & 0 \\ -1 & -1/\sigma & 0 & 0 & 0 & 0 \\ (1-\rho_i)\phi_y & (1-\rho_i)\phi_\pi & -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ i_{t+1} \\ w_{t+1}^s \\ w_{t+1}^d \\ w_{t+1}^i \end{bmatrix} \\
= & \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & 0 & 0 & 0 \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & 0 & 0 & 0 \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & 0 & 0 & 0 \\ -1 & \kappa & 0 & 1 & 0 & 0 \\ -1 & 0 & -1/\sigma & 0 & 1 & 0 \\ 0 & 0 & -\rho_i & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \\ i_t \\ w_t^s \\ w_t^d \\ w_t^i \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1}^y \\ \varepsilon_{t+1}^\pi \\ \varepsilon_{t+1}^i \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1/\sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{t+1}^{ws} \\ \eta_{t+1}^{wd} \\ \eta_{t+1}^{wi} \end{bmatrix}. \tag{14}
\end{aligned}$$

A larger lag length P would cause the system to include more lags of the observables or of the wedges as auxiliary variables, but the same basic idea holds. Instead of trying to estimate (1) directly by deriving what it implies for (6), it is possible to back out the law of motion of w_t by backing (1) out from (6), since (6) is much easier to estimate.

4 Model evaluation using the estimated wedge process

As del Negro and Schorfheide (2009) and Cúrdia and Reis (2011) point out in a Bayesian context, significant off-diagonal elements in the VAR coefficients and in the contemporaneous covariance matrix of the driving process (1) may give a clue as to the dimensions along which a model is misspecified. An understanding of the ways in which a model is misspecified can make it possible to build a better model. Here, I develop that heuristic into a set of tests for overidentifying restrictions, using the New Keynesian model as a pedagogical device.

4.1 The empirical behavior of the wedges

First, I discuss a baseline model ("Model 0") whose parameters are set to $\sigma = 1$; $\kappa = 5.8252$ (which corresponds with a price stickiness coefficient of 0.75 and a labor supply elasticity of one); $\beta = 0.99$; $\phi_\pi = 1.5$; $\phi_y = 0.5/4$; and $\rho_i = 0.75$. I use the OLS estimates for Φ and Σ_z , based on a four-lag VAR with real GDP, GDP deflator inflation, and the three-month Treasury interest rate for the United States from the first quarter of 1985 through the fourth quarter of 2007. This period represents something like a stable monetary regime, and it omits the Great Recession. All variables are detrended using a HP filter with a smoothing parameter of 100,000 to remove the lowest-frequency trends, and the inflation rate and interest rate are expressed as quarterly log gross rates. I filter the data in order to avoid having to model nonstationarity in these three variables. A more serious analysis (as opposed to a pedagogical exercise) would take the issues of first moments, filtering, and nonstationarity more explicitly into account in the context of a more fully-developed model.

The solid lines in Figure 1 show the dynamic cross-correlations among the observables and wedges. The interest rate wedges w_t^i show a particularly strong negative correlation with current and lagged inflation π_t as shown by the subfigure located in the second row and sixth column. There is also a strong contemporaneous correlation between w_t^i and w_t^s . The wedge-extraction exercise has demonstrated that the Taylor rule wedge w_t^i is a particularly poor candidate for a structural monetary policy shock in Model 0—there are large variations in that wedge which seem to be related to past inflation, even allowing for interest rate smoothing.

To understand how the Taylor rule might be misspecified, first I express the Taylor rule wedge as a function of observables. The wedge w_t^i equals $i_t - (1 - \rho_i)\phi_\pi\pi_t - (1 - \rho_i)\phi_y y_t - \rho_i i_{t-1}$. It appears that the calibrated Taylor rule might claim that the contemporaneous response of interest rates to current inflation is much stronger than it is in the data, while it might claim that the response to lagged inflation is weaker than it is in the data. Based on this finding, I now add lags of inflation into the Taylor rule and reduce the dependence of interest

rates on current inflation. The modified Taylor rule allows for the Fed to respond to up to four lags of inflation, which accounts for the fact that it takes time for the Fed to decide upon and implement its new policies as well as to observe inflation data. The modified rule follows the form:

$$i_t = (1 - \rho_i) \sum_{p=1}^4 \phi_{\pi,p} \pi_{t-p+1} + (1 - \rho_i) \phi_y y_t + \rho_i i_{t-1} + w_t^i. \quad (15)$$

I set $\phi_{\pi,1} = \phi_{\pi,2} = \phi_{\pi,3} = \phi_{\pi,4} = 1.5/4$ and redo the previous analysis, calling the new model "Model 1". The resulting cross-correlations among observables and wedges are given by the dashed lines in Figure 1. The interest rate wedge w_t^i behaves much better, especially with respect to lagged inflation. There is no longer a nearly -1 contemporaneous correlation between the interest rate wedge and current inflation. By attenuating the response of interest rates to current inflation, I have found a model of interest rate determination which visually appears to be somewhat less misspecified than the baseline along this particular dimension.

There still seems to be an excessively strong positive relationship between the demand wedge w_t^d and the Taylor rule wedge w_t^i . This suggests that it may be desirable to increase the response of interest rates to output, given by ϕ_y . In "Model 2", I take Model 1 and feed in a value of $\phi_y = 1/4$, which is double its previous value. Figure 2 compares the cross-correlations implied by Model 1 with those implied by Model 2. Model 2 seems to imply less of a relationship than Model 1 between the Taylor rule wedge w_t^i and the other two wedges. It does seem to imply a slightly stronger relationship between the Taylor rule wedge and lagged output than before. When modeling Fed behavior, it may or may not make a certain amount of sense to use a larger value for the response coefficient ϕ_y than that which is typically used in the literature. In general, the wedges approach has indicated that the Fed's response to lagged inflation and its stronger-than-expected response to cyclical

conditions may be important things to take into account when quantitatively evaluating the three-equation New Keynesian model.

4.2 A model misspecification statistic M

The analysis of the New Keynesian model has so far been undertaken heuristically, based on the belief that if wedges are truly structural, they should be uncorrelated with each other or with certain observables. It is possible to formalize this notion and to develop tools to measure the evidence in favor of model misspecification. A model θ is said to be properly specified along dimensions 1 through N when it satisfies N moment restrictions stacked into an N by 1 vector of the form:

$$g(\theta; \bar{\Phi}) = 0, \tag{16}$$

where $\bar{\Phi}$ consists of all of the parameters governing the data generating process for the observables $Z = \{z_t\}_{t=1}^T$. In the context of a VAR, this would equal a column vector of the stacked coefficients $\{\Phi, \Lambda\}$ where Λ consists of the lower triangular elements derived from the Cholesky decomposition of Σ_z such that $\Sigma_z = \Lambda\Lambda'$. The elements of $g(\theta; \bar{\Phi})$ might include the contemporaneous covariances between the wedges, the covariances between current and lagged other wedges, and so on. Since true moments cannot be observed, one must construct the sample counterpart $g(\theta; \tilde{\Phi}(Z))$ of the theoretical moment $g(\theta; \bar{\Phi})$, given an estimate $\tilde{\Phi}(Z)$ of $\bar{\Phi}$, and weigh the evidence against the hypothesis that $g(\theta; \bar{\Phi})$ equals zero.

There is already a large literature which has dealt with evaluating statements like (16). From the perspective of GMM estimation (which includes maximum likelihood as a special case), (16) is a set of overidentifying restrictions. It is possible to transform (16) into a coherent statistic summarizing the evidence in favor of model misspecification along dimensions which the econometrician thinks is important. The traditional approach is to define

an ideal misspecification statistic (which I call M) such that:

$$M(\theta; Z) = [g(\theta; \tilde{\Phi}(Z))]' W(Z) [g(\theta; \tilde{\Phi}(Z))], \quad (17)$$

for some weighting matrix $W(Z)$. If $g(\theta; \tilde{\Phi}(Z))$ has a covariance matrix $S_g(\theta; \tilde{\Phi}(Z))$ such that $S_g(\theta; \tilde{\Phi}(Z)) = E[g(\theta; \tilde{\Phi}(Z))[g(\theta; \tilde{\Phi}(Z))']$, then the most common approach from the GMM literature has been to set $W = [\hat{S}_g(\theta; \tilde{\Phi}(Z))]^{-1}$ for some consistent estimator of $S_g(\theta; \tilde{\Phi}(Z))$ given by $\hat{S}_g(\theta; \tilde{\Phi}(Z))$.

Letting $D(Z)$ equal $\partial g(\theta; \tilde{\Phi}(Z))/\partial \tilde{\Phi}(Z)'$ evaluated at $\tilde{\Phi}(Z)$, and letting $\hat{S}_{\tilde{\Phi}}(Z)$ equal a consistent estimate of $S_{\tilde{\Phi}}$, where $S_{\tilde{\Phi}} = E[(\tilde{\Phi}(Z) - \bar{\Phi})(\tilde{\Phi}(Z) - \bar{\Phi})']$, then the delta method gives one formula for a feasible version of M , such that:

$$M(\theta; Z) = [g(\theta; \tilde{\Phi}(Z))]' [D(\theta; \tilde{\Phi}(Z))' \hat{S}_{\tilde{\Phi}}(Z) D(\theta; \tilde{\Phi}(Z))]^{-1} [g(\theta; \tilde{\Phi}(Z))]/N. \quad (18)$$

In practice, the matrix of derivatives $D(\theta; \tilde{\Phi}(Z))$ may have to be found numerically.

The model misspecification statistic $M(\theta; Z)$ has a number of good properties. It is invariant to the scaling of g . Furthermore, Hansen (1982) shows that under a wide variety of conditions, $g(\theta; \tilde{\Phi}(Z))$ is asymptotically multivariate normal, and so under the null hypothesis (16),

$$N[M(\theta; Z)] \xrightarrow{d} \chi_N^2. \quad (19)$$

It is therefore relatively simple to test the null hypothesis that the model implied by θ is properly specified along the dimensions given by g . One has to first estimate the system (6) using the data Z in order to come up with the estimates $\tilde{\Phi}(Z)$ and $\hat{S}_{\tilde{\Phi}}(Z)$. One can then find $g(\theta; \tilde{\Phi}(Z))$, and then find $D(\theta; \tilde{\Phi}(Z))$, from which one could calculate M based on (18). Furthermore, the magnitude of M is informative in its own right; it has a mean of one under the null hypothesis that (16) is true, and it neither grows nor shrinks with N .

Numbers far above one indicate strong evidence in favor of model misspecification along the dimensions given in (16).

Table 1 contains a complete comparison of Models 0, 1, and 2 based on the misspecification statistic M given in (18), along with the p-values given by the chi-squared test statistic NM given in (19). For the sake of clarity, Appendix A contains pseudocode describing the exact calculations which I have undertaken. I take as the elements of g the model-implied off-diagonal covariances of the wedges at a number of lags. I treat strong relationships among the different wedges as evidence that the baseline model is misspecified. To be precise, I let $g(\theta; \bar{\Phi})$ consist of the set $\{Ew_t^j w_{t-p}^k; j \neq k\}_{p=0}^P$, for values of P ranging from 0 to 4. A low value of P puts most of its weight on contemporaneous correlations while a high value of P puts more of its weight on medium-frequency movements in the wedges.

For four of the five lag windows which I have explored, Model 1 outperforms Model 0 by a wide margin. At lag 0, Model 0 fits the data rather poorly, with an M value of 9.61. Model 1 fits the data at lag 0 much better, with an M value of 2.27. Model 2 appears to fit the data even better than Model 1 with an M value of 1.52. The visual improvements in specification moving between these models shows up in the M statistic as well, when one looks only at contemporaneous correlations.

At lag 1, Model 0 improves drastically, while Model 1 deteriorates slightly and Model 2 deteriorates drastically. The relative ranking between Model 1 and Model 2 reverses. Some aspects of Model 1 fit the lagged data better—in particular, the correlation between the current interest rate shock and the future supply shock is smaller under Model 1 than Model 2. The pattern which is apparent at lag 1 continues all the way through lag 4, with Model 0 performing better than the other models beginning at lag 4. Interestingly, the Taylor rule in Model 1 ends at the third lag, so the reversal in performance between Model 1 and Model 0 at lag 4 provides evidence that the specification of Model 1 can be improved yet further.

The comparison of M statistics can be formalized by borrowing some tools from the SEM literature. Models 1 and 2 seem to be particularly close to each other in their M

statistics when looking at test statistics for lags 1 through 3, and it may be worthwhile to know how precisely the differences misspecification statistics between Model 1 and Model 2 are estimated.

4.3 A relative misspecification statistic R

It is relatively simple to construct a relative misspecification test statistic (denoted by R), which compares the strength of the evidence against the proper specification of θ_1 along some dimensions with the strength of the evidence against the proper specification of θ_2 along the same dimensions, for given parameter sets θ_1 and θ_2 . Such a test would compare $M(\theta_1; Z)$ and $M(\theta_2; Z)$ by taking their relative ratio $R(\theta_1, \theta_2; Z) = M(\theta_1; Z)/M(\theta_2; Z)$. If $Pr(M(\theta_1; Z) > M(\theta_2; Z))$ is large given the estimated process which governs Z , this is evidence that θ_1 appears to be a worse model specification than θ_2 along the dimensions captured by g . If, by contrast, $Pr(M(\theta_1; Z) < M(\theta_2; Z))$ is large over the measure of possible Z , this is evidence that θ_1 is a better model specification than θ_2 along those dimensions. In practice, it is necessary to simulate the distributions of $M(\theta_1; Z)$ and $M(\theta_2; Z)$ for a large number of draws of $\tilde{\Phi}(Z)$ since these quantities will be highly correlated with each other and their ratio will therefore not necessarily follow a standard F distribution. $R(\theta_1, \theta_2; Z)$ is in fact a monotonic transformation of the Normed Fit Index (NFI) of Bentler and Bonett (1980) which is used in the SEM literature. The NFI does not penalize models for additional complexity, which may be an issue when comparing estimated models with different numbers of parameters. Other fit indices contain corrections for degrees of freedom. Yuan (2005) further discusses the role of different fit indices in the SEM literature and their use in evaluating statements about misspecification. In general, these fit indices follow from the types of test statistics discussed in the previous section.

The lower portion of Table 1 shows R statistics which compare Models 0, 1, and 2 each in a pairwise manner with each other, using a zero-lag window and a one-lag window. At windows covering both zero lags and one lag, both Model 1 and Model 2 appear to decisively

show less evidence in favor of misspecification than Model 0. Their R statistics against Model 0 are well below one, and more extreme R statistics are highly improbable given the estimated process governing Z . It is not completely clear that Model 2 dominates Model 1 at a lag of zero. At the MLE of $\tilde{\Phi}(Z)$, there is about an 83% chance of observing an R statistic comparing Model 2 against Model 1 which is larger than the observed statistic of 0.670. By contrast, Model 1 appears to clearly dominate Model 2 at a lag of one, along the dimensions specified by g . While taking lagged inflation into account in the Taylor rule reduces the evidence in favor of the hypothesis that the New Keynesian model is misspecified when looking at lagged data, strengthening the Taylor rule's response to the output gap does not have robust implications for this particular choice of model misspecification statistics.

Despite having abstracted from the issue of identification which is a problem when the wedge process follows a sufficiently general law of motion, is possible to coherently discuss the strength of the evidence on which model specifications exhibit stronger or weaker evidence of misspecification. To the extent that the estimated process governing a DSGE model may violate the commonly-held view that shocks should be orthogonal to each other, such violations contain information about how the model is misspecified along with ways to improve the model.

5 Conclusion

It is possible to quickly and accurately estimate an unrestricted VAR driving process (i.e., a wedge process) for the model errors of a DSGE model, given that the data also follow a VAR process. It is possible to test for overidentifying restrictions on the wedge process and to investigate model misspecification, based on the heuristic idea that some combinations of parameters governing the wedge process are more or less believable than others. Using the example of the Taylor rule in a three-equation new Keynesian model, based on data from the Great Moderation period, I test to see the degree to which off-diagonal covariances

in the wedge process are reflected in three different model specifications. Along several key dimensions, all three of versions of the model appear to be misspecified, though the simplest textbook version of the model shows much more evidence of misspecification than those with modified Taylor rules. It appears that allowing for the Fed to lagged as well as current inflation allows a calibrated Taylor rule to better match the data along my chosen dimensions. Increasing the parameter ϕ_y , which represents Fed's sensitivity to the output gap, results in marginal improvements when looking at contemporaneous data, but it results in stronger evidence of misspecification when looking at lagged data.

Despite the lack of a full identification and estimation strategy, there are coherent ways to discuss which models have more unbelievable implications than others. A chi-squared misspecification statistic (M) can document the strength of the evidence against the null hypothesis that a particular model is correctly specified along a given set of dimensions. A relative misspecification statistic (R) can document the relative strength of the evidence that one model is misspecified in comparison with another. To the extent that one does not truly believe the assumptions governing the structural wedges in DSGE models, one can still talk intelligently about models which appear to be better-specified or worse-specified, and one can then use evidence from the wedge process to build better models.

This paper has deliberately abstracted from the issue of estimation, focusing instead on the issue of testing for model misspecification conditional on a known value of θ . This paper has treated misspecification as a qualitative issue related to model structure rather than as a purely quantitative issue related to parameter values. Posterior predictive checks as discussed by An and Schorfheide (2007) in a Bayesian setting may also be used to flag misspecification using similar logic to that expressed here. The DSGE-VAR literature as discussed by del Negro and Schorfheide (2004, 2009) offers another method to combine the issues of estimation and misspecification into a coherent analysis. Further work on combining the issues of estimation and misspecification from a "wedge" perspective might offer additional guidance as how to balance these issues.

A Pseudocode for generating M and R statistics

This appendix discusses the steps needed to generate the M and R statistics. The generation of these statistics is fairly straightforward once the data are collected, read in, and filtered.

(1) First I estimate the VAR coefficients $\hat{\Phi}$ and $\hat{\Sigma}_z$ on the data Z . I then take $\hat{\Lambda} = \text{chol}(\hat{\Sigma}_z)'$ and stack $\hat{\Phi}$ and the unique elements of $\hat{\Lambda}$ into $\tilde{\Phi}$.

(2) For $\hat{S}_{\tilde{\Phi}}$, I use the quasi-MLE consistent covariance matrix formula given by Hamilton (1994, equation 5.8.7). This covariance matrix is a function of the outer-product covariance matrix of $\tilde{\Phi}$ and the one based on the Hessian of the quasi-likelihood, assuming a homoskedastic Gaussian process for the data. The choice of asymptotic covariance matrices is up to the researcher; I choose this approach because it is more robust than strict Gaussian maximum likelihood to slight misspecification in the likelihood function of the VAR which governs the data.

(3) I evaluate the moment condition $g(\theta_i; \tilde{\Phi}(Z))$ for $i = \{0, 1, 2\}$, where $g(\theta_i; \tilde{\Phi}(Z))$ is an enumeration of the off-diagonal model-implied covariances which I wish to penalize. These covariances are enumerated in Section 4.2. This requires solving the model using equations (7) and (8) from the main text, given the law of motion estimated in step 1, and then deriving the model-consistent covariances.

(4) I then obtain $D(\theta_i; \tilde{\Phi}(Z))$ for $i = \{0, 1, 2\}$ by numerically differentiating $g(\theta_i; \tilde{\Phi}(Z))$ with respect to the elements of $\tilde{\Phi}(Z)$, evaluated at $\tilde{\Phi}(Z)$. This requires repeating step (3) for a large number of slightly different values of $\tilde{\Phi}(Z)$ and then evaluating how $g(\theta_i; \tilde{\Phi}(Z))$ changes when one varies $\tilde{\Phi}(Z)$. I use a small step size for $\tilde{\Phi}(Z)$, in this case 0.001.

(5) Based on the output of the previous steps, I construct $M(\theta_i; Z)$ for $i = \{0, 1, 2\}$ as in (18) and calculate the approximate probability with which $M(\theta_i; Z)$ exceeds its observed value under the null hypothesis of no misspecification, using the chi-squared distribution given in (19). An F test is also reasonable under this circumstance, but I use a large-sample chi-squared approximation instead.

(6) Based on the output of the previous steps, I construct $R(\theta_i, \theta_j; Z)$ for $[i, j] = \{0, 1, 2\}^2$

and test the degree to which $R(\theta_i, \theta_j; Z)$ is statistically distinguishable from one. I simulate the distribution of the ratio of two nonindependent chi-squares in order to generate p-values. I simulate 1,000,000 values for the pair $\{M(\theta_i; Z), M(\theta_j; Z)\}$ by first drawing δ from a multivariate normal distribution with a mean of zero and a covariance matrix given by $\hat{S}_{\tilde{\Phi}}(Z)$. Then I approximate $g(\theta_i; \tilde{\Phi}(Z) + \delta)$ by $g(\theta_i; \tilde{\Phi}(Z)) + D(\theta_i; \tilde{\Phi}(Z))' \delta$, and I approximate $g(\theta_j; \tilde{\Phi}(Z) + \delta)$ the same way. Then I compare the resulting M statistics. If $M(\theta_i; Z)$ is greater than $M(\theta_j; Z)$ I set an indicator variable equal to one; otherwise I set it to zero. I repeat ten million times and take the average of that indicator variable as the p-value in Table 1. A high p-value is an indication that the evidence that model i is misspecified is stronger than the evidence that model j is misspecified.

(7) To generate Figures 1 and 2, I solve for the model using the algorithm discussed in Section 2.2 from the paper using the OLS estimates for the VAR coefficients governing z_t and then simulate that model for one million quarters, tossing out the first 10,000 quarters. Alternately, it is possible to derive expressions for the correlation coefficients of the system by solving the augmented model using equations (7) and (8) from the paper and then expressing the correlation coefficients of the model as a transformation of model's law of motion.

Matlab programs which implement these steps are available from the author upon request.

References

- An, Sungbae, and Frank Schorfheide, 2007. "Bayesian Analysis of DSGE Models". *Econometric Reviews* 26(2-4), pages 113-172.
- Bentler, Peter M., & Bonett, Douglas G., 1980. "Significance tests and goodness-of-fit in the analysis of covariance structures". *Psychological Bulletin* 88(3) 588-600.
- Chari, Varadarajan V., Patrick Kehoe, and Ellen McGrattan, 2007. "Business Cycle Accounting". *Econometrica* 75(3), pages 781-836.
- Cúrdia, Vasco, and Ricardo Reis, 2011. "Correlated Disturbances and U.S. Business Cycles". Manuscript, Columbia University.
- del Negro, Marco, and Frank Schorfheide, 2004. "Priors from General Equilibrium Models for VARs". *International Economic Review* 45(2), pages 643-673.
- del Negro, Marco, and Frank Schorfheide, 2009. "Monetary Policy Analysis with Potentially Misspecified Models". *American Economic Review* 99(4), pages 1415-1450.
- Hamilton, James D., 1994. *Time Series Analysis*. Princeton University Press.
- Hansen, Lars Peter, 1982. "Large Sample Properties of Generalized Method of Moments Estimators". *Econometrica* 50(4), pages 1029-1054.
- Ravenna, Federico, 2007. "Vector autoregressions and reduced form representations of DSGE models". *Journal of Monetary Economics* 54(7), pages 2048-2064.
- Reicher, Christopher Phillip, 2013. "A note on the identification of dynamic economic models with generalized shock processes." Kiel Working Paper 1821.
- Sims, Christopher A., 2002. "Solving Linear Rational Expectations Models". *Computational Economics* 20(1-2), pages 1-20.

Smets, Frank, and Rafael Wouters, 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". *American Economic Review* 97(3), pages 586–606.

Šustek, Roman, 2011. "Monetary Business Cycle Accounting." *Review of Economic Dynamics* 14(4), pages 592-612.

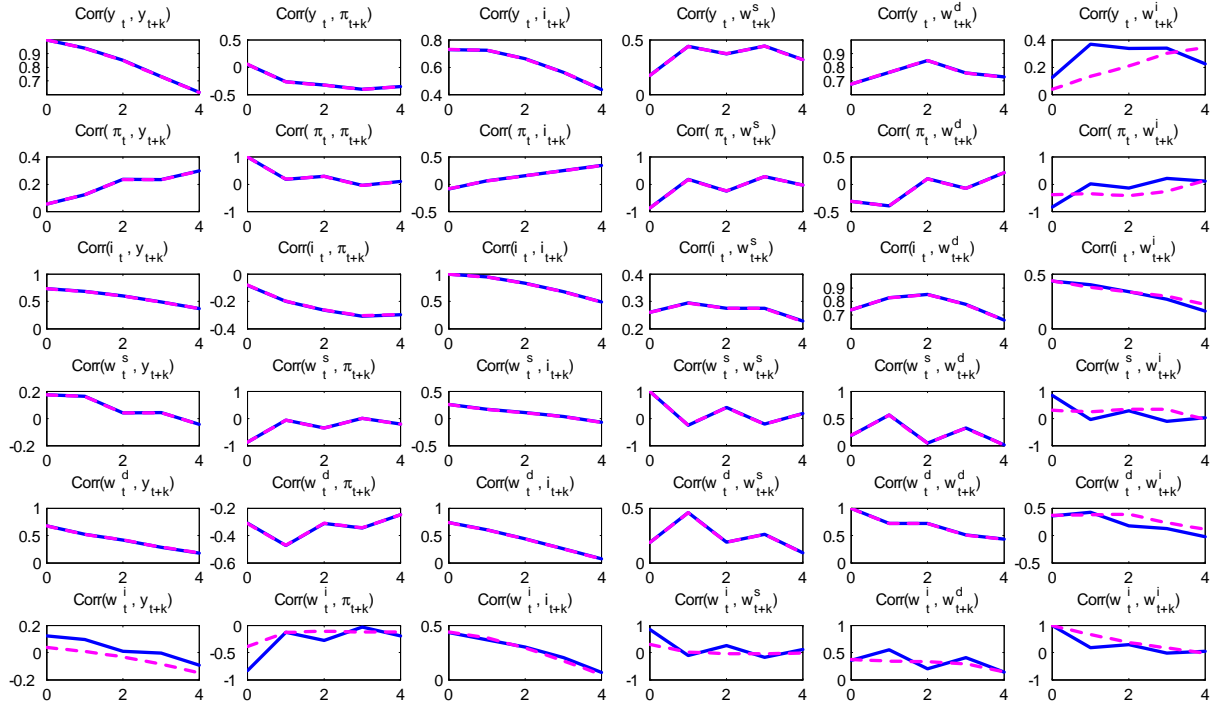
Yuan, Ke-Hai, 2005. "Fit Indices Versus Test Statistics". *Multivariate Behavioral Research* 40(1), pages 115-148.

Table 1: Comparison of Models 0, 1, and 2

Parameter / Statistic	Model 0	Model 1	Model 2
Taylor rule coefficient $\phi_{\pi,1}$	1.5	1.5/4	1.5/4
Taylor rule coefficients $\phi_{\pi,2}$, $\phi_{\pi,3}$, and $\phi_{\pi,4}$	0	1.5/4	1.5/4
Taylor rule coefficient ϕ_y	0.5/4	0.5/4	1.0/4
Taylor rule coefficient ρ_i	0.75	0.75	0.75
Misspecification statistic $M(\theta; Z)$, 0 lags	9.61	2.27	1.52
Misspecification test p-value	<0.001	0.078	0.206
Misspecification statistic $M(\theta; Z)$, 1 lag	5.64	2.74	3.36
Misspecification test p-value	<0.001	0.003	<0.001
Misspecification statistic $M(\theta; Z)$, 2 lags	4.14	3.49	3.62
Misspecification test p-value	<0.001	<0.001	<0.001
Misspecification statistic $M(\theta; Z)$, 3 lags	3.56	2.97	3.05
Misspecification test p-value	<0.001	<0.001	<0.001
Misspecification statistic $M(\theta; Z)$, 4 lags	3.11	3.20	3.29
Misspecification test p-value	<0.001	<0.001	<0.001
Comparison with Model 0 $R(\theta_x, \theta_0; Z)$, 0 lags	1	0.240	0.160
$Pr(M(\theta_x; Z) < M(\theta_0; Z))$	-	>0.999	0.999
Comparison with Model 1 $R(\theta_x, \theta_1; Z)$, 0 lags	4.24	1	0.670
$Pr(M(\theta_x; Z) < M(\theta_1; Z))$	<0.001	-	0.833
Comparison with Model 2 $R(\theta_x, \theta_2; Z)$, 0 lags	6.31	1.49	1
$Pr(M(\theta_x; Z) < M(\theta_2; Z))$	<0.001	0.167	-
Comparison with Model 0 $R(\theta_x, \theta_0; Z)$, 1 lag	1	0.49	0.60
$Pr(M(\theta_x; Z) < M(\theta_0; Z))$	-	>0.999	0.999
Comparison with Model 1 $R(\theta_x, \theta_1; Z)$, 1 lag	2.06	1	1.23
$Pr(M(\theta_x; Z) < M(\theta_1; Z))$	<0.001	-	0.002
Comparison with Model 2 $R(\theta_x, \theta_2; Z)$, 1 lag	1.68	0.82	1
$Pr(M(\theta_x; Z) < M(\theta_2; Z))$	0.001	0.998	-

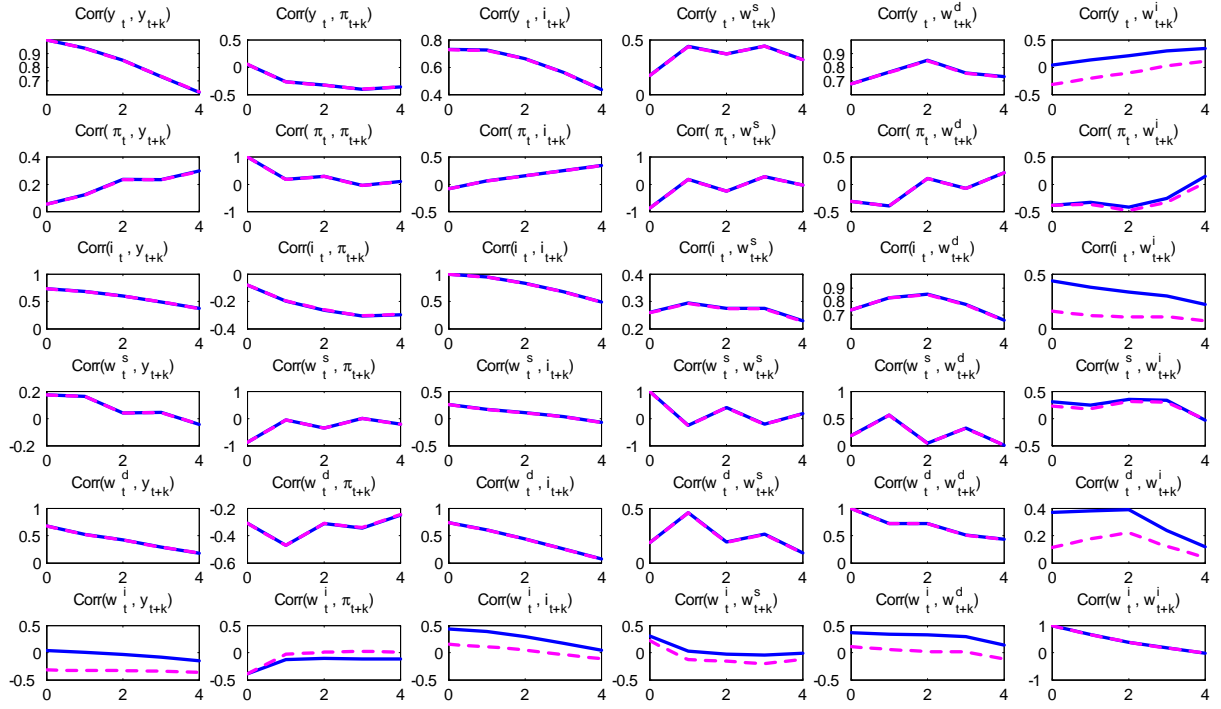
This table compares Models 0, 1, and 2. The top portion of this table show the calibrated parameter values in each model specification. The middle portion of this table shows the M statistics for each model and the p-value of a test of non-misspecification for wedge cross-covariances of lags 0 through 4. The bottom portion of this table shows the R statistics for each model pair and the p-value of a test of equal misspecification between the models. A low p-value indicates that the model in each of the three columns appears to be more misspecified than the comparison model along the dimensions covered by g .

Figure 1: Cross-correlations under Model 0 and Model 1



This figure compares the cross-correlations among the data and the wedges for the three-equation New Keynesian model with a Taylor rule which responds only to contemporaneous inflation (Model 0: $\phi_{\pi,1} = 1.5$, solid blue lines) versus a specification where the Fed responds to inflation over the past year (Model 1: $\phi_{\pi,1} = \phi_{\pi,2} = \phi_{\pi,3} = \phi_{\pi,4} = 1.5/4$, dashed magenta lines). Time is on the x axis and is in quarters. Correlation coefficients are on the y axis. The wedges were estimated using U.S. data from 1985 to 2007, using four lags of HP filtered output, inflation, and interest rates.

Figure 2: Cross-correlations under Model 1 and Model 2



This figure compares the cross-correlations among the data and the wedges for the three-equation New Keynesian model with a Taylor rule which responds weakly to output (Model 1: $\phi_y = 0.5/4$, solid blue lines) versus a specification where the Fed responds more strongly to output (Model 2: $\phi_y = 1/4$, dashed magenta lines). Both models allow for the Fed to respond to the past year's inflation with coefficients of $\phi_{\pi,1} = \phi_{\pi,2} = \phi_{\pi,3} = \phi_{\pi,4} = 1.5/4$. Time is on the x axis and is in quarters. Correlation coefficients are on the y axis. The wedges were estimated using U.S. data from 1985 to 2007, using four lags of HP filtered output, inflation, and interest rates.