SURPRISING COMPARATIVE PROPERTIES OF MONETARY MODELS: RESULTS FROM A NEW MODEL DATABASE

John B. Taylor and Volker Wieland*

Abstract—In this paper, we investigate the comparative properties of empirically estimated monetary models of the U.S. economy using a new database of models designed for such investigations. We focus on three representative models due to Christiano, Eichenbaum, and Evans (2005), Smets and Wouters (2007), and Taylor (1993a). Although these models differ in terms of structure, estimation method, sample period, and data vintage, we find surprisingly similar economic impacts of unanticipated changes in the federal funds rate. However, optimized monetary policy rules differ across models and lack robustness. Model averaging offers an effective strategy for improving the robustness of policy rules.

I. Introduction

EVER since the 1970s revolution in macroeconomics, monetary economists have been building quantitative models that incorporate the fundamental ideas of the Lucas critique, time inconsistency, and forward-looking expectations in order to evaluate monetary policy more effectively. The common characteristic of these monetary models, compared with earlier models, is the combination of rational expectations, staggered price and wage setting, and policy rules, all of which have proved essential to policy evaluation.

Over the years, the number of monetary models with these characteristics has grown rapidly as the ideas have been applied in more countries, as researchers have endeavored to improve on existing models by building new ones, and as more data shed light on the monetary transmission process. The last decade, in particular, has witnessed a surge of macroeconomic model building as researchers have further developed the microeconomic foundations of monetary models and applied new estimation methods. In our view, it is important for research progress to document and compare these models and assess the value of model improvements in terms of the objectives of monetary policy evaluation. Keeping track of the different models is also important for monetary policy in practice because by checking the robustness of policy in different models, one can better assess policy.

Received for publication May 25, 2009. Revision accepted for publication December 17, 2010.

* Taylor: Stanford University and Hoover Institution; Wieland: University of Frankfurt.

V.W. is grateful for financial support by a Willem Duisenberg Research Fellowship at the European Central Bank while this research project was initiated. Excellent research assistance was provided by Elena Afanasyeva, Tobias Cwik, and Maik Wolters from Goethe University Frankfurt. V.W. acknowledges research assistance funding from European Community grant MONFISPOL under grant agreement SSH-CT-2009-225149. Comments by Mark Watson, Frank Smets, three anonymous referees, and session participants at the AEA Meetings 2009, the German Economic Association Monetary Committee Meeting 2009, and the NBER Summer Institute 2009 Monetary Economics Group also proved very helpful. All remaining errors are our own.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00220.

With these model comparison and robustness goals in mind, we have recently created a new monetary model database, an interactive collection of models that can be simulated, optimized, and compared. The database can be used for model comparison projects and policy robustness exercises. Perhaps because of the large number of models and the time and cost of bringing modelers together, there have not been many model comparison projects and robustness exercises in recent years. In fact the most recent policy robustness exercise, which we both participated in, occurred ten years ago as part of an NBER conference.[1]

Our monetary model database provides a new platform that makes model comparison much easier than in the past and allows individual researchers easy access to a wide variety of macroeconomic models and a standard set of relevant benchmarks.[2] We hope in particular that many central banks will participate and benefit from this effort as a means of getting feedback on model development efforts.

This paper investigates the implications of three well-known models in the model database for monetary policy in the U.S. economy. The first model, which is a multicountry model of the G-7 economies built more than fifteen years ago, has been used extensively in the earlier model comparison projects. It is described in detail in Taylor (1993a). The other two models are the best-known representatives of the most recent generation of empirically estimated new Keynesian models: the Christiano, Eichenbaum, and Evans (2005) model of the United States and the Smets and Wouters (2007) model of the United States.

The latter two models incorporate the most recent methodological advances in terms of modeling the implications of optimizing behavior of households and firms. They also utilize new estimation methods. The Christiano et al. (2005) model is estimated to fit the dynamic responses of key macroeconomic variables to a monetary policy shock identified with a structural vector autoregression. The Smets and Wouters (2007) model is estimated with Bayesian methods to fit the dynamic properties of a range of key variables in response to a full set of shocks.

[1] The results are reported in Taylor (1999). Several of the models in this earlier comparison and robustness exercise are also included in our new monetary model database, including Rotemberg-Woodford (1999), McCallum and Nelson (1999), and Taylor (1993a).

[2] See the appendix of this paper for the list of 38 models and Wieland et al. (2009) for a detailed exposition of the platform for model comparison. The model base includes small calibrated textbook-style models, estimated medium- and large-scale models of the U.S. and euro-area economies, and some estimated open economy and multicountry models. Software and models are available for download from http://www.macromodelbase.com. This platform relies on the DYNARE software for model solution and may be used with Matlab. For further information on DYNARE, see Collard and Juillard (2001) and Juillard (1996) and http://www.cepremap.cnrs.fr/dynare/dynare.

The Review of Economics and Statistics, August 2012, 94(3): 800–816

First, we examine and compare the monetary transmission process in each model by studying the impact of monetary policy shocks in each. Second, we calculate and compare the optimal monetary policy rules within a certain simple class for each of the models. Third, we evaluate the robustness of these policy rules by examining their effects in each of the other models relative to the rule that would be optimal for the respective model.

The model comparison and robustness analysis reveals some surprising results. Even though the two more recent models differ from the Taylor (1993a) model in terms of economic structure, estimation method, data sample, and data vintage, they imply almost identical estimates of the response of U.S. GDP to an unexpected change in the federal funds rate, that is, to a monetary policy shock. This result is particularly surprising in light of earlier findings by Levin, Wieland, and Williams (1999, 2003) indicating that a number of models built after Taylor (1993a) exhibit quite different estimates of the impact of a monetary policy shock and the monetary transmission mechanism.[3] We also compare the dynamic responses to other shocks. Interestingly, the impact of the main financial shock, that is, the risk premium shock, on U.S. GDP is also quite similar in the Smets and Wouters (2007) and the Taylor (1993a) models. This finding is of interest in light of the dramatic increase in risk premia observed since the start of the financial crisis in August 2007.[4] Differences emerge with regard to the consequences of other demand and supply shocks.

The analysis of optimized simple interest rate rules reveals further interesting similarities and differences across the three models. All three models prefer rules that include the lagged interest rate in addition to inflation deviations from target and output deviations from potential. The two more recent new Keynesian models favor the inclusion of the growth rate of output gaps.

The robustness exercise, however, delivers more nuanced results. Model-specific rules with interest rate smoothing and output gaps are not robust. Some degree of robustness can be recovered by focusing on two-parameter rules with inflation and the output gap or three-parameter rules with interest rate smoothing, inflation, and the deviation of output growth from trend instead of output gap growth. This increase in robustness in relation to other models comes at the cost of significant performance deterioration in the original model. Fortunately, however, model comparison offers an avenue for improving over the robustness properties of model-specific rules. Rules that are optimized with respect to the average loss across multiple models achieve very good robustness properties at much lower cost.

[3] For example, the model of Fuhrer and Moore (1995) and the Federal Reserve's FRB/US model of Reifschneider et al. (1999) both exhibited longer-lasting effects of policy shocks on U.S. GDP that peak several quarters later than in Taylor (1993a). See Levin et al. (1999, 2003) for a comparison.

[4] As Smets and Wouters (2007) noted, the risk premium shock represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households and has similar effects as so-called net worth shocks in models with an explicit financial sector such as Bernanke et al. (1999).

II. Brief Description of the Models

A. *Taylor (1993a)*

This is an econometrically estimated rational expectations model fit to data from the G7 economies for the period 1971:1 to 1986:4. All of our simulations focus on the United States. The model was built to evaluate monetary policy rules and was used in the original design of the Taylor rule. It has also been part of several model comparison exercises, including Bryant et al. (1989), Klein (1991), Bryant, Hooper, and Mann (1993), and Taylor (1999). Shiller (1991) compared this model to the "old Keynesian" models of the prerational expectations era, and he found large differences in the impact of monetary policy due largely to the assumptions of rational expectations and more structural models of wage and price stickiness.

To model wage and price stickiness, Taylor (1993a) used the staggered wage and price setting approach rather than ad hoc lags of prices or wages that characterized the older prerational expectations models. However, because the Taylor (1993a) model was empirically estimated, it used neither the simple example of constant-length four-quarter contracts presented in Taylor (1980) nor the geometrically distributed contract weights proposed by Calvo (1983). Rather it lets the weights have a general distribution, which is empirically estimated using aggregate wage data in the different countries. In Japan, some synchronization is allowed for.

The financial sector is based on several no-arbitrage conditions for the term structure of interest rates and the exchange rate. Expectations of future interest rates affect consumption and investment, and exchange rates affect net exports. Slow adjustment of consumption and investment is explained by adjustment costs such as habit formation or accelerator dynamics. A core principle of this model is that after a monetary shock, the economy returns to a growth trend, which is assumed to be exogenous to monetary policy as in the classical dichotomy.

Most of the equations of the model were estimated with Hansen's instrumental variables estimation method, with the exception of the staggered wage-setting equations, which were estimated with maximum likelihood.

B. *Christiano, Eichenbaum, and Evans (2005)*

Many of the equations in the model of Christiano et al. (CEE, 2005, in the following) exhibit similarities to the equations in the Taylor model, but they are explicitly derived log-linear approximations of the first-order conditions of optimizing representative firms and households. Their model also assumes staggered contracts but with Calvo weights and backward-looking indexation in those periods when prices and wages are not set optimally. Long-

run growth and short-run fluctuations are modeled jointly rather than separately as in Taylor's model. Thus, the CEE (2005) model explicitly accounts for labor supply dynamics as well as the interaction of investment demand, capital accumulation, and utilization. Furthermore, their model includes a cost channel of monetary policy. Firms must borrow working capital to finance their wage bill. Thus, monetary policy rates have an immediate impact on firms' profitability.

The CEE (2005) model was estimated for the U.S. economy over the period 1959:2–2001:4 by matching the impulse response function to the monetary shock in a structural vector autoregression (VAR). An important assumption of the VAR that carries over to the model is that monetary policy innovations affect the interest rate in the same quarter but other variables, including output and inflation, only by the following quarter.

The monetary policy innovation represents the single, exogenous economic shock in the original CEE model. However, additional shocks can be incorporated in the structural model, and the variance of such shocks may be estimated using the same methodology. The additional shocks would first be identified in the structural VAR. Then the parameters of the structural model including innovation variances would be reestimated by matching the impulse response functions implied by the model with their empirical counterparts from the VAR. Altig et al. (2004; ACEL, 2004 in the following), follow this approach and identify two additional shocks: a neutral and an investment-specific technology shock. These shocks exhibit serial correlation and have permanent effects on the level of productivity. Together with the monetary policy shock, they account for about 50% of the variation in output. The impulse response function for the monetary policy shock in ACEL (2004) is almost identical to CEE (2005). Therefore, we use the ACEL (2004) parameterization of the CEE model for the computational analysis in our paper. A drawback of this model is that it does not yet provide a complete characterization of the observed output and inflation volatility.

The CEE model, which was initially circulated in 2001, represented the first medium-sized, estimated example of the new generation of New Keynesian dynamic stochastic general equilibrium models explicitly derived from optimizing behavior of representative households and firms.[5] It stimulated the development of similar optimization-based models for many other countries once Smets and Wouters (2003) showed how to make use of new advances in Bayesian techniques (see Geweke, 1999, and Schorfheide, 2000) in estimating such models.

C. *Smets and Wouters (2007)*

The model of the U.S. economy estimated by Smets and Wouters (2007; SW, 2007, in the following) with U.S. data from 1966:1 to 2004:4 may be viewed as an extended version of the CEE/ACEL model. The SW model contains a greater set of macroeconomic shocks and aims to fully explain the variation in key variables, such as aggregate output and its components, as well as inflation, wages, and interest rates. They use a Bayesian estimation methodology that allows the use of priors on model parameters informed from theory and literature. The posterior distributions then incorporate the information in the available macroeconomic data. Whenever the data do not help in pinpointing parameter values precisely, theoretical priors dominate. Such priors can in some cases be based on evidence from microeconomic studies. The Bayesian estimation methodology has quickly been popularized and widely applied by researchers in central banks and academia. It has been implemented for use with the DYNARE software that we also utilize in our model base.

Smets and Wouters (2007) modify some of the structural assumptions embodied in the CEE/ACEL model. In the long run, the SW model is consistent with a balanced steady-state growth path driven by deterministic labor-augmenting technological progress. While the CEE model assumes wages and prices are indexed to the last period's inflation rate in the absence of a Calvo-style signal, the SW model allows firms to index to a weighted average of lagged and steady-state inflation. Furthermore, SW drop two more assumptions that have important short-run implications in the CEE/ACEL model. First, they do not impose the delayed effect of monetary policy on other variables that CEE built into the structural model so as to match the constraints required by the structural VAR to identify the monetary policy shock. Second, they do not require firms to borrow working capital to pay the wage bill. Thus, the so-called cost channel is absent from the model. Smets and Wouters note that they did not find this channel necessary for fitting the dynamics in U.S. data. In our simulations, we will also investigate the implications of adopting the SW assumptions of no cost channel and no timing constraints on monetary policy shocks in the original CEE/ACEL model.

III. Shocks to Monetary Policy as Deviations from Two Policy Rules

We first use the model database to assess the extent of differences between models regarding the transmission of monetary policy to output and inflation. To this end, we compare the effect of monetary policy shocks in the three models. A monetary policy shock is defined as a surprise deviation from systematic policy behavior that is characterized by interest rate policy rules.

In our comparison, we focus on two estimated rules used by SW (2007) and CEE (2005), respectively, to characterize systematic central bank policy. Smets and Wouters estimate the coefficients of this interest rate rule along with the other equations in their model. We refer to it as the SW rule in

[5] The paper was published in 2001 as NBER working paper no. 8403.

the remainder of the paper. They call it a generalized Taylor rule because it includes the lagged federal funds rate, the lagged output gap, and a serially correlated policy shock, in addition to the current inflation rate and output gap that appear in the original Taylor (1993b) rule. The SW rule implies the following setting for the federal funds rate, i_t:

$$i_t = 0.81i_{t-1} + 0.39\pi^a_{t+1} + 0.97y_t - 0.90y_{t-1} + \varepsilon^i_t,$$
$$\text{where} \quad \varepsilon^i_t = 0.15\varepsilon^i_{t-1} + \eta^i_t. \tag{1}$$

Here, π^a_t-refers to the annualized, quarterly inflation rate and y_t to the output gap.[6] In the SW and CEE model, the gap measure used in the policy rule is defined as the difference between the actual output level and the level that would be realized if prices adjust flexibly to macroeconomic shocks—the so-called flex-price output level.[7] In the Taylor model (and the original Taylor rule), the output gap is defined as the difference between actual output and long-run potential output as measured by the trend. The policy shock is denoted by ε^i_t and follows a first-order autoregressive process with an independent and identically distributed (i.i.d.) normal error term, η^i_t. As a result of serial correlation and the inclusion of the lagged interest rate in the reaction function, an i.i.d. innovation will have a persistent effect on nominal interest rates and due to price rigidity also on real rates and aggregate output.

CEE (2005) define the central bank's policy rule in terms of a reaction function for the growth rate of money.[8] They identify monetary policy shocks in a structural VAR as orthogonal innovations to the interest rate reaction function. Then they estimate the parameters of the structural model, including the parameters of the money growth rule, by matching the impulse response in the structural model and the VAR. In addition, they contrast their findings under the money growth rule with the effect of a policy shock under an extended Taylor rule for the federal funds rate:[9]

$$i_t = 0.80i_{t-1} + 0.3E_t\pi^a_{t+1} + 0.08y_t + \varepsilon^i_t. \tag{2}$$

Just like the SW rule, it incorporates partial adjustment to the lagged federal funds rate. However, it is forward looking and responds to the expected inflation rate for the upcoming quarter. The coefficient on the output gap is much smaller than in the SW rule and does not include the lag of the output gap. The policy shock is IID. In the following, we refer to this rule as the CEE rule.

IV. Monetary Policy Shocks in Three Monetary Models of the U.S. Economy

We compare the consequences of a monetary policy shock in the Taylor, SW, and CEE/ACEL models to shed light on their implications for the transmission of Federal Reserve interest rate decisions to aggregate output and inflation. In particular, we want to find out to what extent the current-generation DSGE models, CEE/ACEL (2004) and SW (2007), imply quantitatively different effects of monetary policy than the model by Taylor (1993a). Since the models differ in terms of economic structure and parameter estimates are obtained for different data series, estimation periods, and data vintages, we would expect to obtain quantitatively different assessments of the monetary transmission mechanism.

Figure 1 reports the consequences of a 1 percentage point shock to the federal funds rate for nominal interest rates, output, and inflation. The panels on the left-hand side refer to the outcomes when the Federal Reserve sets interest rates following the initial shock according to the prescriptions of the SW rule, while the right-hand-side panels refer to the outcome under the CEE rule. Each panel shows the findings from four model simulations. The thick solid line refers to the SW model, the thin solid line with squares to the TAYLOR model, the thin solid line with filled circles to the CEE/ACEL model, and the dashed line to the CEE/ACEL model with SW assumptions.[10]

Surprisingly, the effect of the policy shock on real output and inflation given a common policy rule is very similar in the four models. For example, under the SW rule, the nominal interest rate increases on impact by 0.8 to 1 percentage points and then returns slowly to steady state, real output falls over three to four quarters to a trough of about -0.35% before returning to steady state, and inflation declines more slowly with a trough of about -20 basis points roughly two to three quarters later than output.

The quantitative implications for real output in the Taylor (1993a) and SW (2007) models are almost identical. The outcome under the CEE/ACEL model initially differs slightly from the other two models. In the quarter of the shock, we observe a tiny increase in output, while inflation does not react at all. From the second quarter onward, output declines to the same extent as in the other two models,

[6] Note that the response coefficients differ from the values reported in SW 2007. In equation (1), interest and inflation rates are annualized, while SW used quarterly rates. The original specification in SW 2007 corresponds to $i^q_t = (1 - 0.81)(2.04\pi^q_t + 0.09y_t) + 0.22\Delta y_t + 0.81i^q_{t-1} + \varepsilon^i_t$, where the superscript q refers to quarterly rates that are not annualized.

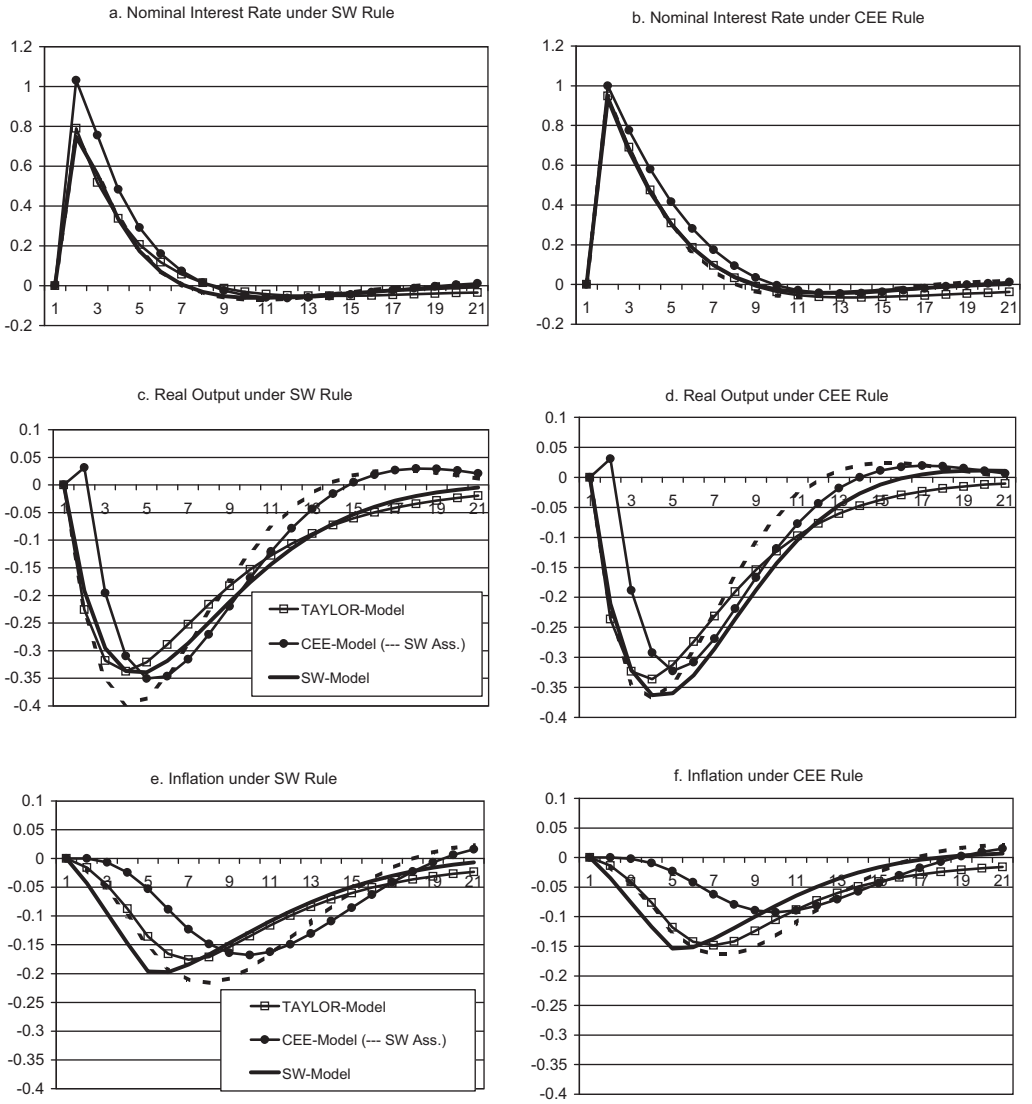
[7] Smets and Wouters set wage and price markup shocks equal to 0 in the derivation of the flex-price output measure used to define their output gap.

[8] CEE (2005) and ACEL (2004) model monetary policy in terms of innovations to the growth rate of money that they denote by μ_t: $\mu_t = \mu + \theta_0\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}...$

[9] Note that we use annualized interest and inflation rates and transcribe the CEE rule accordingly. CEE (2005) define their rule as $i^q_t = (1 - 0.80)(1.5E_t\pi^q_{t+1} + 0.1y_t) + 0.8i^q_{t-1} + \varepsilon^i_t$. They attribute this estimated rule to Clarida et al. (1999). However, the coefficients reported in Clarida et al. (1999) are different. Their rule corresponds to $i_t = (1 - 0.79)(2.15E_t\pi_{t+1} + 0.93y_t) + 0.79i_{t-1} + \varepsilon^i_t$.

[10] The CEE/ACEL model with SW assumptions implies the following modifications. We remove the timing constraints that were imposed on the structural model by the authors so that it coincides with the identification restrictions on the VAR that they used to obtain impulse responses for the monetary policy shock. Furthermore, we remove the constraint from the ACEL model that requires firms to finance the wage bill by borrowing cash in advance from a financial intermediary. As a result of this constraint, the interest rate has a direct effect on firms' costs.

FIGURE 1.—THE EFFECT OF A POLICY SHOCK ON INTEREST RATES, OUTPUT AND INFLATION
1 PERCENTAGE POINT INCREASE IN THE NOMINAL POLICY RATE

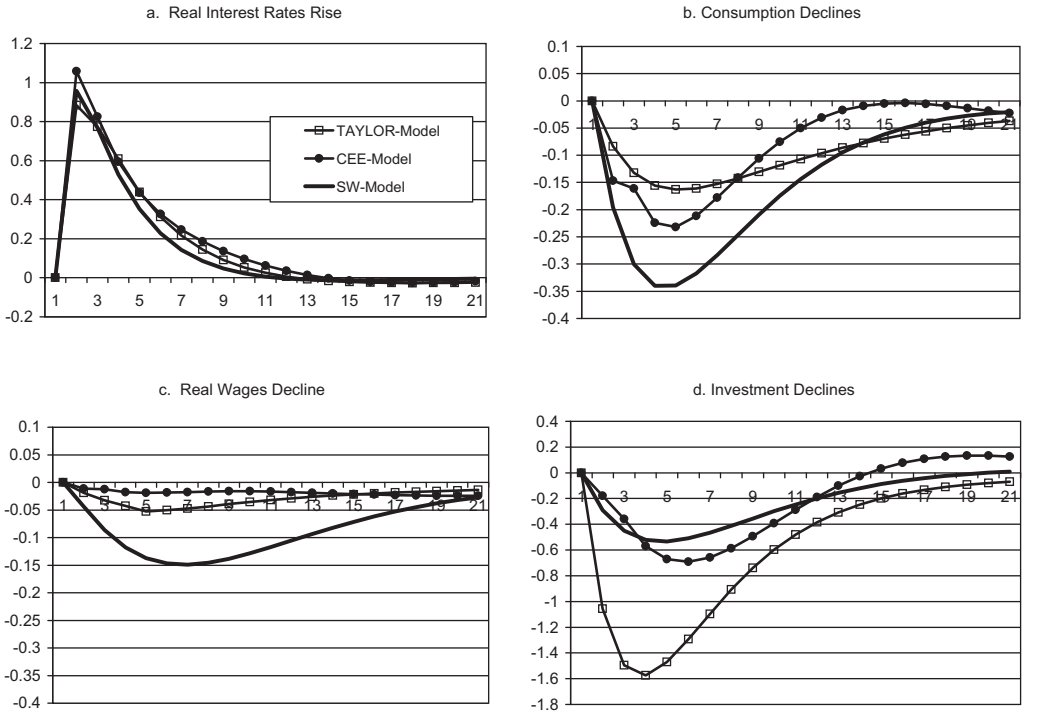


but the profile is shifted roughly one quarter into the future. The decline in inflation is similarly delayed. Once we implement the CEE/ACEL model with the SW assumptions of no timing constraint on policy and no cost channel, the timing of output and inflation dynamics is more similar to the other two models.

The outcome of a monetary policy shock given the Fed follows the CEE rule is shown in the right-hand-side panels of figure 1. Again, the magnitude of the effect of the policy shock on real output and inflation is almost identical in the Taylor model, the SW model, and the ACEL/CEE model, particularly when the last model is implemented with the SW assumptions. Furthermore, the reduction in output is very similar to the case when the Fed follows the SW rule. The decline in inflation is a bit smaller.

The original Lucas critique stated that a change in the systematic component of policy can have important implications for the dynamics of macroeconomic variables. Thus, it is not surprising that the output and inflation effects of monetary policy shocks change if we consider a wider set of monetary policy rules. For example, in the case of the original Taylor (1993b) rule, an i.i.d. policy shock would influence the nominal interest rate for only one period, because the Taylor rule does not include the lagged interest rate. We have investigated the real output effects of a monetary policy shock with different response coefficients (for example, a four times smaller response to output), different inflation measures (such as year-on-year inflation), and different rules, such as the original Taylor rule or the benchmark rules considered in Levin et al. (2003) and Kuester and Wieland (2010). Different rules have quite different implications for the real consequences of monetary policy shocks. However, the Taylor model, the SW model, and the CEE model continue to imply surprisingly similar dynamics of aggregate real output and inflation in response to a policy shock for a given, common policy rule.

FIGURE 2.—COMMON ASPECTS OF THE TRANSMISSION MECHANISM IN THE THREE MODELS (SW RULE)



The finding that the two best-known models of the recent generation of new Keynesian models provide very similar estimates of the impact of a policy shock on U.S. real GDP as the model of Taylor (1993a) is particularly surprising in light of earlier comparison projects. For example, the comparison in Levin et al. (1999, 2003) indicated that models built and estimated after Taylor (1993a), such as the model of Fuhrer and Moore (1995) or the Federal Reserve's FRB/US model of Reifschneider, Tetlow, and Williams (1999), provided different assessments of the U.S. monetary transmission mechanism. In particular, these models suggested that the impact of monetary policy shocks on real output would be longer lasting and reach its peak more than a year after the initial impulse. This view is often considered conventional wisdom among practitioners. The model database associated with this paper also allows users to replicate the impulse response function comparison in the Fuhrer and Moore and FRB/US models.

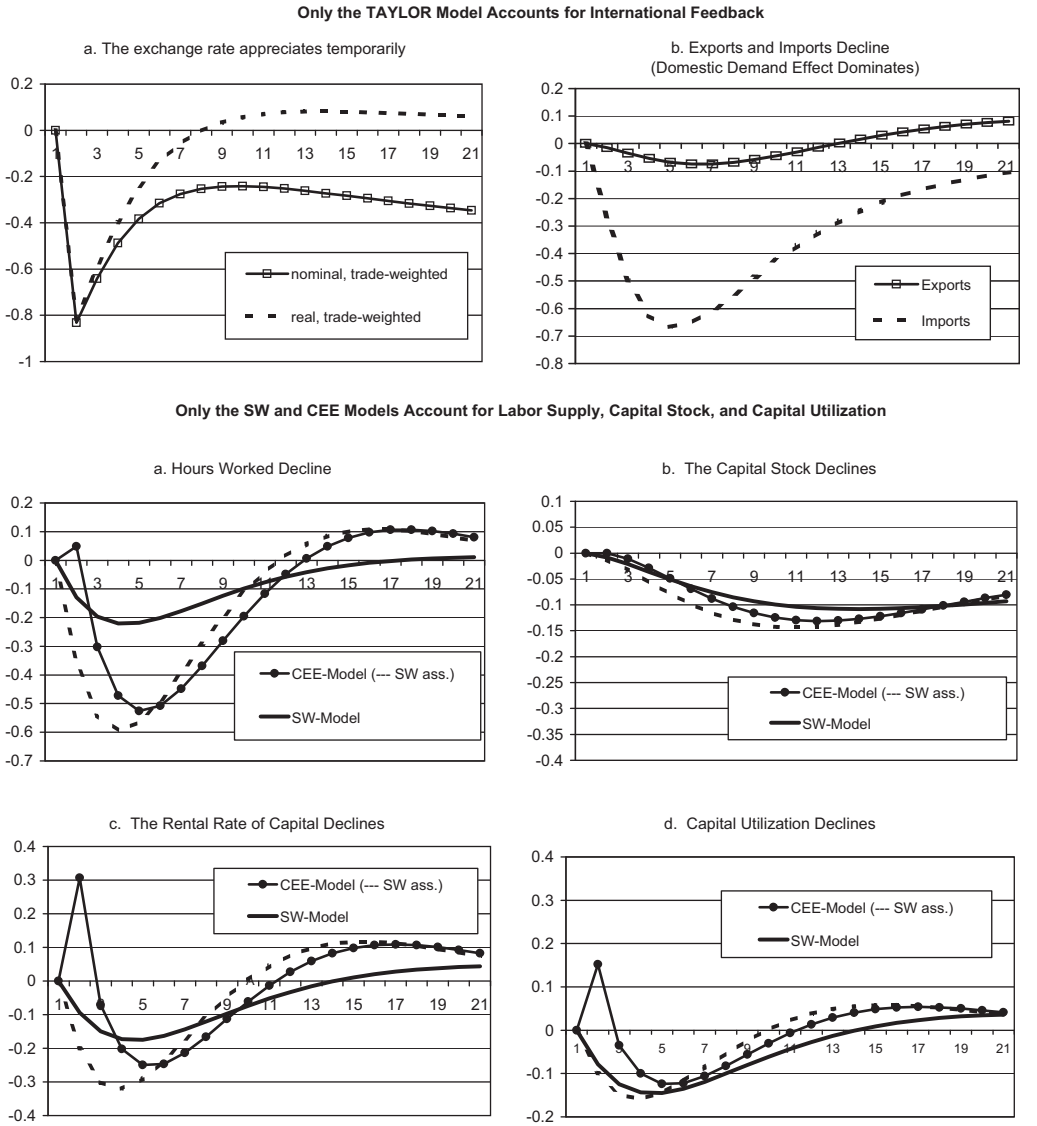
So far we have focused on the overall effect of the policy shock on output and inflation. Now we turn to the effects on other macroeconomic variables. Figure 2 illustrates some additional common aspects of the transmission mechanism in the three models of the U.S. economy, while figure 3 highlights interesting differences. Monetary policy is assumed to follow the SW rule after the policy shock.[11] The real interest rate increases almost to the same extent in all three models as shown in figure 2a. As a result, aggregate consumption and aggregate investment decline. The decline in consumption is smaller in the Taylor model than in the other two models, and the decline in investment is much greater. The quantitative comparison of the dynamics of GDP components, however, is hampered by the fact that the models use different deflators in generating real consumption and investment series.[12] Another similarity regarding monetary policy transmission in the three models is that real wages decline along with aggregate demand following the monetary policy shock.

The three models also exhibit some interesting differences regarding monetary policy transmission. For example, figures 3a and 3b indicate that only the Taylor model accounts for international feedback effects. As a result of the policy shock, the U.S. dollar appreciates temporarily in real trade-weighted terms. Exports and imports both decline. However, the fall in imports is much greater than in exports, and as a result, net exports increase. The strong decline in imports occurs due to the domestic demand effect that figures very importantly in the U.S. import demand equation. The resulting increase in net exports partly offsets the impact of the large negative decline in investment demand on aggregate output in the Taylor model. Furthermore, figures 3c to 3f illustrate that only the SW and CEE models account for the effects of the policy shock on labor supply,

[11] Similar figures for the case of the CEE rule are provided in the online appendix.

[12] While the Taylor model simulates the components of GDP in real terms, the simulations in the SW and CEE models concern the nominal components divided by the GDP deflator. It is not possible to make the series directly comparable because none of the models accounts for the consumption and investment deflators separately from the GDP deflator.

FIGURE 3.—DIFFERENCES IN THE TRANSMISSION MECHANISM IN THE THREE MODELS (SW RULE)



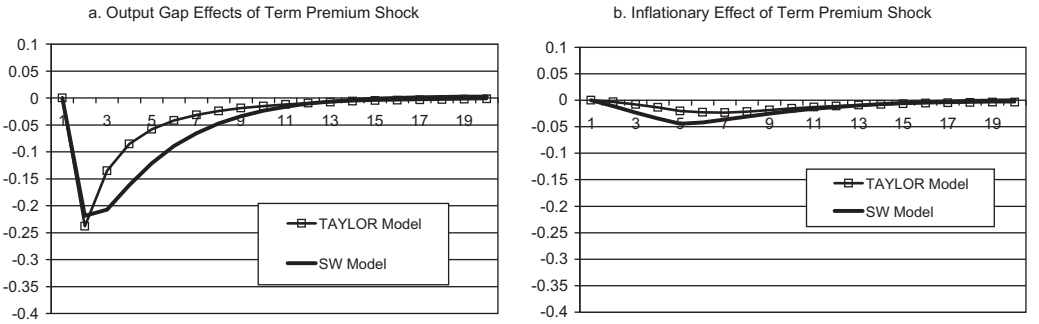
capital stock, the rental rate of capital, and capital utilization. All four measures decline in response to the monetary shock. This explanation of supply-side dynamics is missing from the Taylor model.

V. Other Shocks and Their Implications for Policy Design

Unexpected changes in monetary policy are of interest in order to identify aspects of the monetary transmission mechanism. When it comes to the question of policy design, however, the standard recommendation is to avoid policy surprises since they only generate additional output and inflation volatility. Instead, optimal and robust policy design focuses on the proper choice of the variables and the magnitude of the response coefficients in the policy rule that characterizes the systematic component of monetary policy. The policy rule is then designed to stabilize output and inflation in the event of shocks emanating from other sectors of the economy. In this respect, it is of interest to review and compare the potential sources of economic shocks in the three models under consideration.

In light of the recent financial crisis, we start by comparing the effect of particular financial shocks. Only the Taylor and SW models contain such shocks. Figure 4 illustrates the effect of an increase in the term premium by 1 percentage point on real output and inflation in the Taylor and SW models. The initial impact of these shocks on real output is almost identical in the two models and lies between -0.22 and -0.24% of output. This finding is particularly surprising since the shocks are estimated quite differently in the two models. In the Taylor model, the term premium shock is estimated from the term structure equation directly using data on short- and long-term interest rates, that is, the fed-

FIGURE 4.—TERM PREMIUM SHOCK IN THE TAYLOR AND SW MODELS (SW RULE)
1 PERCENTAGE POINT INCREASE IN THE TERM PREMIUM



eral funds rate versus ten-year U.S. treasuries. In the SW model, the risk premium shock is estimated from the consumption and investment equation. It assumes the term structure relation implicitly but uses no data on long-term rates. In earlier work on the euro area, Smets and Wouters (2003) included instead a consumption demand or preference shock. This shock is omitted in their model of the U.S. economy to keep the number of shocks in line with the number of observed variables. SW emphasize that the premium shock represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households and has similar effects as so-called net worth shocks in models with an explicit financial sector such as Bernanke, Gertler, and Gilchrist (1999).[13]

Figure 5 provides a comparison of what could be termed "demand" or "spending" shocks in the three models. These are shocks that push output and inflation in the same direction. The Taylor model contains many such shocks. Figures 5a and 5b show the effects of shocks to nondurables consumption, equipment investment, inventory investment, government spending, and import demand on the output gap and inflation. The SW model contains two shocks of this type: an exogenous spending shock that comprises government spending as well as net exports and an investment-specific technology shock. The ACEL model contains an investment-specific technology shock that initially lowers inflation but then raises it. It has stronger long-term effects than the investment-specific technology shock in SW (2007).

Figure 6 compares supply shocks in the three models—shocks that push output and inflation in opposite directions. The Taylor model has a number of such shocks, in particular, innovations to the contract wage equations, the final goods price equation, import prices, and export prices. The SW model contains price mark-up and wage mark-up shocks that are somewhat similar to the contract wage and aggregate price shocks in the Taylor model. Only the SW and the ACEL models include neutral technology shocks. In the ACEL model, these shocks have a long-term effect on productivity growth, while their effect on productivity growth in the SW model is temporary.

Comparing the three models, it is important to keep in mind that only the Taylor and SW model aim to fully explain the variation in the macroeconomic variables included in the model as an outcome of exogenous shocks and endogenous propagation. The ACEL model aims only to explain the part of the variation that is caused by the three shocks in the structural VAR that was used to identify them. Figures 5 and 6 indicate that the investment-specific and neutral technology shocks in the ACEL model have negligible effects on inflation. Consequently, the ACEL model omits most sources of inflation volatility outside of policy shocks and is of limited usefulness for designing monetary policy rules. With this caution in mind, we will nevertheless explore the implications of the ACEL model for policy design together with the other two models.

VI. Optimal Simple Policy Rules in the Taylor, CEE/ACEL, and SW Models

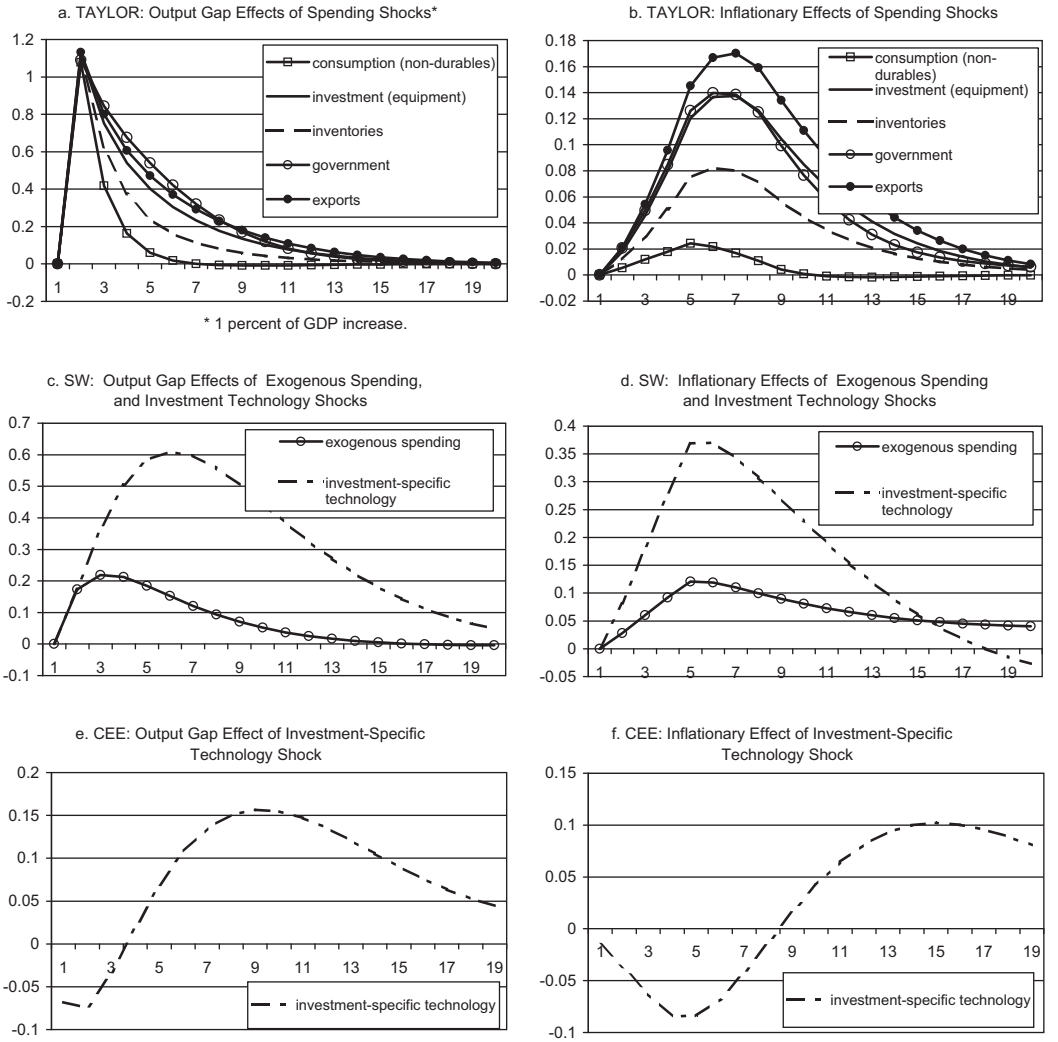
The first question on policy design that we address concerns the models' recommendations for the optimal policy response to a small number of variables in a simple interest rate rule. We start by considering rules that incorporate a policy response to two variables: the current year-on-year inflation rate and the output gap as in the original Taylor (1993b) rule:

$$i_t = \alpha\pi_t + \beta_0 y_t. \tag{3}$$

In the SW and ACEL models, the output gap y is defined as the deviation of actual output from the level of output that would be realized if the price level were fully flexible. This flexible-price output varies in response to some of the eco-

[13] In the model file available from the AER Web site along with SW (2007), the shock is multiplied with the negative value of the consumption elasticity. This is consistent with figure 2 of that paper, where the shock appears as a "demand" shock, that is, an increase has a positive effect on output. It is not consistent with equation (2) in SW (2007) that identifies the shock as a risk premium shock. In this case, an increase has a negative effect. We have modified the model file consistent with the notation as risk premium shock in equation (2) in SW (2007). In addition, we have checked that reestimating the SW model with the shock entering the consumption Euler equation as defined by equation (2) in their paper does not have an important effect on the parameter estimates.

FIGURE 5.—"DEMAND" SHOCKS IN THE TAYLOR, SW, AND CEE MODELS (SW RULE)
1 PERCENT INCREASE IN THE RELEVANT VARIABLES



nomic shocks. We use the same definition of flexible price output as in SW (2007). In the Taylor model, the gap is calculated relative to a measure of potential that grows at an exogenous rate.

In a second step, we extend the rule to include the lagged nominal interest rate as in Levin et al. (1999, 2003):

$$i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t. \tag{4}$$

Then we also include the lagged output gap as in the estimated rule in the SW (2007) model:

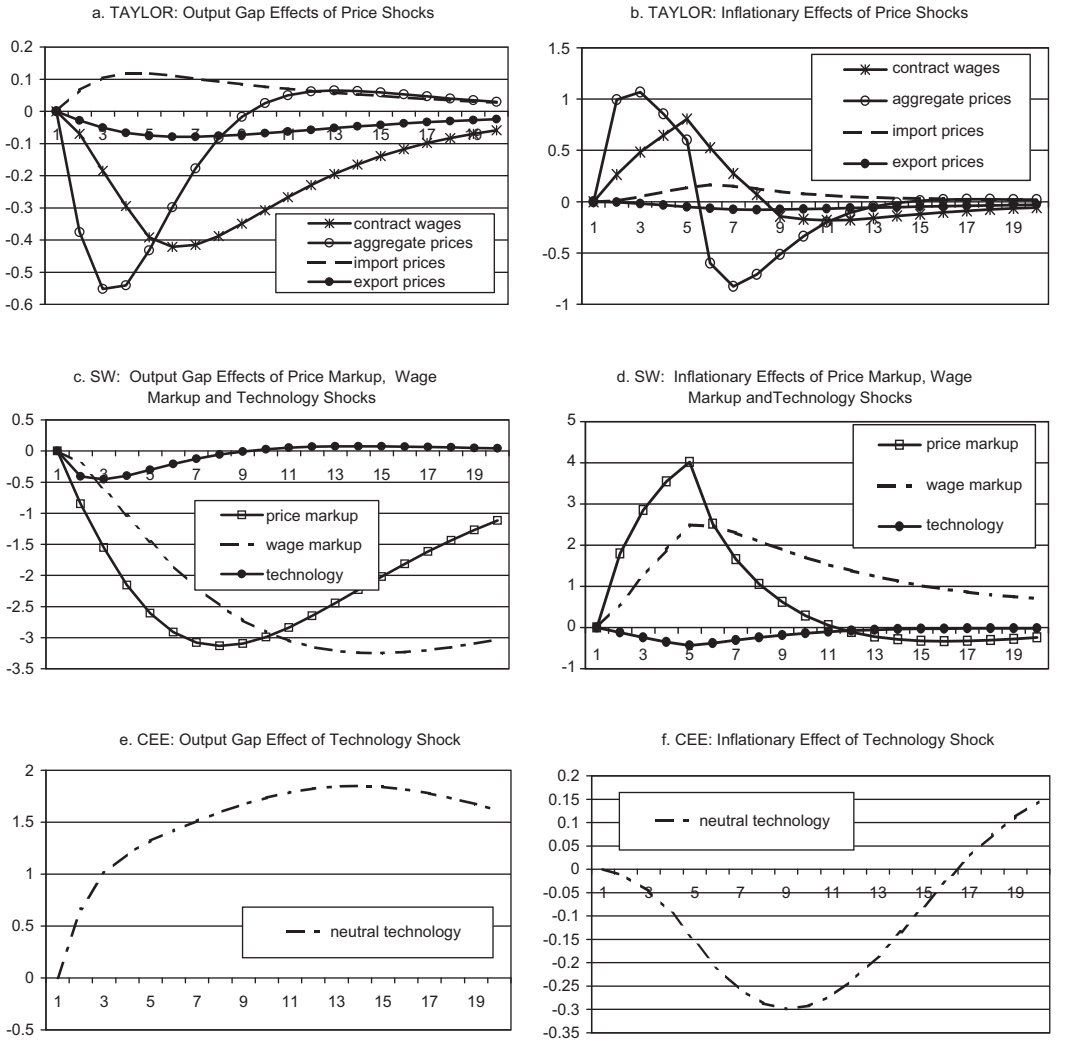
$$i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}. \tag{5}$$

We choose the response coefficients of the rules, that is, (ρ, α, β_0, β_1), in each of the models by minimizing a loss function L that includes the unconditional variances of inflation, the output gap, and the change of the nominal interest rate:

$$L = \mathrm{Var}(\pi) + \lambda_y \mathrm{Var}(y) + \lambda_{\Delta i} \mathrm{Var}(\Delta i). \tag{6}$$

This form of loss function has been used extensively in earlier analyses, including the model comparison studies. With $\lambda_{\Delta i}$=0, it corresponds to the unconditional expectation of a second-order approximation of household utility in a small New Keynesian model derived from microeconomic foundations as shown in Rotemberg and Woodford (1999). The magnitude of the implied value of λ_y is very sensitive to the particular specification of overlapping nominal contracts: random-duration "Calvo-style" contracts imply a very low value on the order of 0.01, whereas fixed-duration "Taylor-style" contracts imply a value near unity (see Erceg & Levin, 2006). For this reason, we consider values of $\lambda_y \in \{0, 0.5, 1\}$. In addition, we assign a positive weight to interest volatility and consider values of $\lambda_{\Delta i} \in \{0.5, 1\}$. It is intended to capture central banks' well-known tendency to smooth interest rates and avoid extreme values of optimized response coefficients that would be very far from empirical

FIGURE 6.—SHORT-RUN AND LONG-RUN "SUPPLY" SHOCKS IN TAYLOR, SW, AND CEE MODELS (SW RULE)
1 PERCENT INCREASE IN THE RELEVANT VARIABLES



observation and regularly violate the nonnegativity constraint on nominal interest rates (see Woodford, 1999).

The optimized response coefficients are shown in table 1. It reports results for two-, three-, and four-parameter rules in the Taylor, SW, and CEE/ACEL models. The central bank's objective is assumed to assign a weight of unity to inflation and interest rate volatility and a weight of either 0 or unity to output gap volatility.[14] First, with regard to two-parameter rules, all three models prescribe a large response coefficient on inflation and a small coefficient on the output gap if the output gap does not appear in the loss function. If the output gap receives equal weight in the loss function, the optimal coefficient on output increases but remains quite a bit below the response to inflation. The coefficient on inflation declines in the SW and CEE/ACEL models but increases in the Taylor model when output appears in the loss function.

For three-parameter rules, the optimized value of the coefficient on the lagged nominal interest rate is near unity. This property applies in all three models and with different values of the objective function weights except for one case that we discuss below. The coefficients on inflation are much smaller than in the two-parameter rules, but they typically remain positive.

In the ACEL model, the loss function is very flat. There appear to be multiple local optima, and the global optimum we identify has very extreme coefficients in the case of the three-parameter rule with a positive weight on output gap volatility in the loss function.[15] As noted earlier, a weakness of the ACEL model is that it contains only two technology shocks that explain little of the variation of inflation and output gaps but have permanent effects on the growth

[14] Additional findings for a weight of 0.5 on the unconditional variance of the change of the nominal interest rate are reported in the appendix available online. Further sensitivity studies for intermediate weights have been conducted and are available from the authors on request.

[15] A local optimum at less extreme values is observed for $\rho = 0.01$, $\alpha = 2.9$, $\beta_0 = 0.5$.

TABLE 1.—OPTIMAL SIMPLE POLICY RULES
RULES: $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1} + \beta_\Delta \Delta y_t$

	Loss($\lambda_y = 0$): Var(π) + Var(Δi)					**Loss($\lambda_y = 1$)**: Var(π) + Var(y) + Var(Δi)				
Rule/Model	ρ	α	β_0	β_1	β_Δ	ρ	α	β_0	β_1	β_Δ
	2 parameters (Gap)[a]					$\alpha\pi_t + \beta_0 y_t$				
Taylor		2.54	0.19				3.00	0.52		
SW		2.33	−0.10				2.04	0.26		
CEE/ACEL		4.45	0.28				2.57	0.45		
	3 parameters (Gap)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$				
Taylor	0.98	0.37	0.09			0.98	0.21	0.53		
SW	1.06	0.49	0.01			1.13	0.012	0.015		
CEE/ACEL	0.97	0.99	0.02			2.84	7.85	−2.12		
	4 parameters (Gaps)					$\rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$				
Taylor	0.98	0.37	0.07	0.02		0.96	0.18	0.41	0.19	
SW	1.06	0.46	−0.03	0.03		1.07	0.16	1.63	−1.62	
CEE/ACEL	1.01	1.11	0.18	−0.18		1.04	0.51	2.24	−2.30	
	3 parameters (Growth)[b]					$\rho i_{t-1} + \alpha\pi_t + \beta_\Delta \Delta y_t$				
Taylor	1.01	0.52			0.07	1.13	0.40			0.68
SW	1.03	0.48			−0.01	1.01	0.20			1.04
CEE/ACEL	1.02	1.07			−0.002	0	3.71			0.002

The loss function includes the variance of inflation and the variance of the first difference of nominal interest rates with a weight of unity, $\lambda_{\Delta i} = 1$. λ_y denotes the weight on the variance of the output gap.

[a] In the Taylor model, the output gap denotes the difference between actual and trend output. In the SW and ACEL models, it is the difference to the level realized under flexible prices given current macroeconomic shocks.

[b] The output growth measure Δy_t is defined relative to steady-state/trend output growth in all three models.

of steady-state output. The ACEL model contains no short-run demand and supply shocks as do the other two models. For this reason, the model may not be considered suitable in its current form for an evaluation of the role of interest rate rules in stabilization policy. Nevertheless, we continue to replicate the analysis conducted in the other two models also in the ACEL model throughout this paper.[16]

Next, we turn to the rules with four parameters that include the lagged output gap in addition to current output, inflation, and the lagged interest rate. The coefficients on the lagged interest rate typically remain near unity. Interestingly, the coefficient on the lagged output gap, that is, β_1, in the CEE/ACEL and SW models is almost equal to $-\beta_0$, the coefficient on the current output gap. Thus, the CEE/ACEL and SW models appear to desire a policy response to the growth rate of the output gap rather than its level. In fact, restricting $\beta_1 = -\beta_0$ and reoptimizing the response coefficients in these models implies a coefficient of 1.65 in the SW and 2.0 in the ACEL model, respectively. Changes in the other response coefficients are very limited. By contrast, in the Taylor model, which uses trend output as a measure of potential, the optimal coefficients on current and lagged output gaps are both positive.

Different findings between the Taylor model and the SW and CEE/ACEL models may be due to different definitions of potential output. The flex-price output level used as a measure of potential in the SW and CEE/ACEL models exhibits substantial variation due to economic shocks, and its growth rate may deviate substantially from trend growth.[17] Thus, simply differencing the output gap in our policy rule does not eliminate the effect of different concepts of potential output on the optimized response coefficients. Instead, we proceed to evaluate the performance of a fourth class of rules that respond to the deviation of actual GDP growth from trend (or steady-state) growth, denoted by Δy_t:

$$i_t = \rho i_{t-1} + \alpha\pi_t + \beta_\Delta \Delta y_t. \quad (7)$$

In this manner, potential output growth is defined similarly across the three models. Researchers such as Orphanides (2003a) have recommended such rules as a way to reduce the impact of central bank misperceptions about the level of potential output on interest rate setting.[18] The last three rows in table 1 report the optimal coefficients of the three-parameter policy rule with deviations of actual from steady-state output growth.

If output gap variability does not appear in the loss function, ($\lambda_y = 0$), the optimal coefficient on output growth, $\beta_{\Delta y}$, is very close to 0, just as in the three-parameter rules with the output gap. If output variability receives a weight of unity in the loss function, the optimal interest rate rule responds positively to output growth, at least in the Taylor and SW models. In the ACEL models, it is near 0. Thus, in the SW and ACEL models, it matters quite a lot whether

[16] Following the suggestion of an anonymous referee, we have investigated whether the SW model exhibits similar properties as the ACEL model if the number of shocks is reduced to the investment-specific and the neutral technology shock as in ACEL. We find that the response coefficients on inflation and the output gap in the two-parameter and three-parameter rules increase in absolute terms. However, the three-parameter rule in the SW model does not take the extreme coefficient values observed in the ACEL model, nor do we observe multiple local optima as in the ACEL model. We make these findings available along with other material in the online appendix.

[17] A number of recent contributions have emphasized the differences between flex price measures of potential and more traditional views on the trending components of real activity (see Palmqvist, 2007; Basu & Fernald, 2009; and Gupta, 2009).

[18] See also Schmitt-Grohe and Uribe (2004) and Burriel, Fernández-Villaverde, and Rubio-Ramírez (2009). Beck and Wieland (2008) have instead proposed a nonlinear cross-checking mechanism that would correct the prescriptions from an output gap-based rule whenever there is statistical evidence of distorted policy outcomes, but take advantage of gap estimates in normal times.

TABLE 2.—INCREASE IN LOSS WHEN REDUCING THE NUMBER OF PARAMETERS IN THE RULE
PERCENTAGE INCREASE (INCREASE IN IIP)

Models	Four vs. Three Parameters (Gaps)	Four Parameters (Gaps) vs. Three Parameters (Growth)	Three vs. Two Parameters (Gaps)
		Loss($\lambda_y = 0$): $\mathrm{Var}(\pi) + \mathrm{Var}(\Delta i)$	
Taylor	0.12% (0.001)	13.5% (0.10)	278% (1.38)
SW	0.22% (0.001)	1.40% (0.01)	316% (0.78)
CEE/ACEL	5.10% (0.001)	10.0% (0.003)	229% (0.04)
		Loss($\lambda_y = 1$): $\mathrm{Var}(\pi) + \mathrm{Var}(y) + \mathrm{Var}(\Delta i)$	
Taylor	1.81% (0.07)	67.1% (1.61)	98.8% (2.14)
SW	10.6% (0.47)	18.1% (0.76)	25.6% (1.17)
CEE/ACEL	14.4% (0.11)	36.7% (0.22)	9.67% (0.11)

The values in parentheses measure the increase in absolute loss in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

the rule uses the deviation of actual GDP growth from trend growth or from flexible-price output growth.

Table 2 reports on the relative stabilization performance with two-, three-, and four-parameter rules. Two different measures are reported: the percentage increase in loss and, in parentheses, the absolute increase in loss when one reduces the number of parameters (and therefore variables) in the policy rule starting from the case of four-parameter rules. In the following, we focus on the absolute loss differences because the percentage differences tend to give misleading signals.

The particular measure of the increase in absolute loss that is shown is the implied inflation variability premium proposed by Kuester and Wieland (2010) (referred to as the IIP in the following). This measure translates a particular increase in absolute loss into the increase in the standard deviation of inflation (in percentage point terms) that would raise the loss to the same extent keeping all else equal (for a constant output or interest volatility). The advantage of this measure is that it is easily interpreted in practical terms and therefore provides a clear signal of those properties of interest rate rules that are of economic importance.

To give an example, consider the number in the fourth row and third column of table 2 in parentheses. Its value is 2.14, and it implies the following: if the Taylor model represents the U.S. economy and the central bank considers using the optimized two-parameter rule instead of an optimized three-parameter rule, and if the central bank's loss function assigns equal weight to output and inflation, the resulting increase in loss (due to higher inflation, output and interest volatility) is equivalent to an increase in the standard deviation of inflation of 2.14 percentage points, all else equal. This difference is economically important. Although it is the largest IIP reported in the table, the associated percentage increase of 98.8% is only the fourth largest in the table. The third-largest percentage increase in the table is 229%. It is associated with a switch from the three-parameter to the two-parameter rule in the ACEL model when the central bank's loss function assigns 0 weight to output volatility. However, the associated IIP of 0.04 is tiny. Thus, the particular switch in rule is economically irrelevant in spite of the large percentage increase in loss. In this case, the reason is that the ACEL model contains only two shocks that cause little inflation volatility and very small losses.

The findings in table 2 suggest little additional benefit from including the lagged output gap in the rule. Dropping the lagged output gap from the rule barely increases the central bank's loss. The associated IIPs in the first column of table 2 lie between 0.001 and 0.47. However, it appears very beneficial to include the lagged interest rate in the rule. Dropping the lagged interest rate from the rule and moving from three to two response parameters implies an economically significant increase in the central bank's loss function, in particular, in the SW and Taylor models, where it is equivalent to an increase in the standard deviation of inflation by 1 and 2 percentage points, respectively (third column in table 2). Among three-parameter rules, the rule with the output gap performs better than the rule with the growth rate of output (in deviation from trend growth) across all three models. As shown in the middle column of table 2, the IIPs relative to the four-parameter rule are uniformly greater for the growth rate than the gap version. They are particularly large in the Taylor model. However, the growth rate version of the three-parameter rule still performs better than the two-parameter rule with inflation and the output gap.

VII. Robustness

What if the model used by the central bank in designing a policy rule is not a good representation of the economy and one of the other two models provides a much better representation of the U.S. economy? In other words, how robust are model-specific optimized policy rules with respect to the range of model uncertainty reflected in the three models considered in this paper? Table 3 provides answers to these questions. Robustness is measured in the following manner. The rule optimized for model X is implemented in model Y. The resulting loss in model Y is compared to the loss that would be realized under the rule with the same number of parameters that has been optimized for that particular model. The difference is expressed in terms of IIP only.

TABLE 3.—ROBUSTNESS OF POLICY RULES
INCREASE IN IIP WHEN A RULE OPTIMIZED IN MODEL X IS USED IN MODEL Y AND EVALUATED RELATIVE TO THE SAME TYPE OF RULE OPTIMIZED IN MODEL Y

IIP if Evaluated in Model:	Two Parameters	Three Parameters (Gap)	Three Parameters (Growth)	Four Parameters (Gaps)
		Loss(λ_y=0): $\mathrm{Var}(\pi) + \mathrm{Var}(\Delta i)$		
Rules Optimized in Taylor Model[a]				
SW	0.37	0.83	0.01	0.90
ACEL	0.03	0.12	0.01	0.14
Rules Optimized in SW Model				
Taylor	0.27	0.13	0.03	0.15
ACEL	0.15	0.02	0.01	0.02
Rules Optimized in ACEL Model				
SW	0.54	0.11	0.10	0.09
Taylor	0.76	0.27	0.25	0.34
		Loss(λ_y=1): $\mathrm{Var}(\pi) + \mathrm{Var}(y) + \mathrm{Var}(\Delta i)$		
Rules Optimized in Taylor Model				
SW	0.17	5.41	0.66	7.18
ACEL	0.001	M.E.[b]	0.31	M.E.
Rules Optimized in SW Model				
Taylor	0.86	3.20	1.05	2.71
ACEL	0.03	0.21	0.44	0.13
Rules Optimized in ACEL Model				
SW	0.07	108	1.69	0.53
Taylor	0.12	24.9	1.40	3.85

The values in this table concern the increase in absolute loss in model Y under a rule optimized for model X relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

[a]Rules: 2 parameters: $i_t = \alpha\pi_t + \beta_0 y_t$; 3 parameters (Gap): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t$; 3 parameters(Growth): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_\Delta \Delta y_t$; 4 parameters (Gaps): $i_t = \rho i_{t-1} + \alpha\pi_t + \beta_0 y_t + \beta_1 y_{t-1}$.

[b]M.E. refers to indeterminacy and the existence of multiple self-fulfilling equilibria.

The findings in table 3 show that from the perspective of a central bank that aims to minimize inflation and interest rate volatility but assigns no weight to output volatility ($\lambda_y = 0$), all four classes of policy rules are quite robust. Typically a rule optimized in one of the models performs quite well in any of the other models compared to the best possible rule with the same number of parameters in that model.

Unfortunately, the preceding conclusion is almost completely reversed when one takes the perspective of a policymaker who cares equally about output and inflation volatility, that is, when $\lambda_y = 1$. In this case, only the two-parameter rules remain fairly robust. The lack of robustness is most pronounced for three- and four-parameter rules that use output gaps. While these rules offer substantial performance improvements when the true model is known, performance can deteriorate markedly if the economy is better approximated by another model. For example, using the four-parameter rule that is optimal in the SW model instead in the Taylor model implies an IIP of 2.71. Alternatively, the four-parameter rule optimized for the Taylor model implies an IIP of 7.18 in the SW model and generates multiple equilibria in the ACEL model.

Similar problems arise with regard to three-parameter rules that use the output gap, even if the CEE/ACEL model is excluded from the robustness analysis because of its odd behavior under such rules as discussed earlier. As shown in the second column of table 3, the rule optimized in the Taylor model implies an IIP of 5.41 in the SW model, while the rule optimized for the SW model delivers an IIP of 3.20 in the Taylor model. Replacing the output gap in the three-parameter rules with the deviation of output growth from its trend improves their robustness properties at the cost of substantial performance deterioration in the true model as shown previously in table 2. However, the IIPs are not negligible and remain near or above unity in three cases, two of which concern the rule optimized in the ACEL model.

Only the rules with two parameters that respond to inflation and the current output gap deliver a fairly robust stabilization performance across the three models. The IIPs are always substantially below unity and often near 0. Thus, a policymaker with a strong preference for robustness against model uncertainty might prefer to choose an optimized two-parameter rule that responds to inflation and the output gap but not the lagged interest rate.

Unfortunately, such rules perform quite a bit worse than rules with interest rate smoothing when it is known which of the models best captures the true dynamics in the economy. To quantify this loss, we reevaluate robustness with respect to the best four-parameter rule when the model is known rather than the best rule of the same class. With respect to this benchmark, two-parameter rules exhibit IIPs of 2.64 (SW rule in Taylor model) and 1.53 (Taylor rule in SW model), respectively. Thus, they remain more robust than three- and four-parameter rules with output gaps. However, three-parameter rules that replace the output gap with the deviation of actual GDP growth from trend perform slightly better from this perspective as long as the ACEL model is excluded from the comparison. They exhibit IIPs of 2.28 (SW rule in Taylor model) and 1.21 (Taylor rule in SW model), respectively, when compared to the four-parameter rule optimized for the correct model.

Using the model database, however, it is possible to produce policy recommendations that are more robust than

TABLE 4.—OPTIMIZED MODEL-AVERAGING RULES
OBJECTIVE: Min $\sum_{m\in M}\frac{1}{3}(\mathrm{Var}(\pi_m)+\mathrm{Var}(y_m)+\mathrm{Var}(\Delta i_m))$;
RULES: $i_t=\rho i_{t-1}+\alpha\pi_t+\beta_0 y_t+\beta_1 y_{t-1}+\beta_\Delta \Delta y_t$

Set of Equally Weighted Models: $M=\{SW, TAYLOR, ACEL\}$	ρ	α	β_0	β_1	β_Δ
2-parameter rule (gap)		2.75	0.52		
3-parameter rule (gap)	1.05	0.41	0.23		
3-parameter rule (growth)	1.09	0.20			0.76
4-parameter rule (gap)	1.06	0.19	0.67	−0.59	

those based on a single model. For example, one may optimize a particular policy rule with respect to multiple models by minimizing the average loss across models. This approach has been proposed by Levin et al. (2003) and Brock, Durlauf, and West (2003), among others. In this case, the response coefficients of the rules, (ρ, α, β_0, β_1, β_Δ), are chosen to minimize the average loss across the three models:

$$\sum_{m\in M}\frac{1}{3}L_m=\sum_{m\in M}\frac{1}{3}\left(\mathrm{Var}(\pi_m)+\lambda_y\mathrm{Var}(y_m)+\lambda_{\Delta i}\mathrm{Var}(\Delta i_m)\right). \tag{8}$$

Here, the subscript m refers to a particular element of $M=\{TAYLOR, SW, ACEL\}$—the set of available models. We focus on the performance of such rules in those cases where model-specific rules were not robust, that is, when the central bank assigns similar weights to output and inflation in the loss function. The parameter values for the model averaging rules are reported in table 4. The two-parameter rules remain fairly similar to the model-specific optimization because those were already quite robust. The interest-smoothing coefficient for three- and four-parameter rules now lies very close to unity, between the values that are optimal in the SW and the Taylor models. The response to inflation is small but positive, ranging from 0.2 to 0.4, depending on whether the rules include current and lagged output gaps or the deviation of output growth from trend. Response coefficients on the current output gap, output gap, growth or output growth deviations from trend vary between 0.2 and 0.8.

As shown in table 5, model averaging generally improves the robustness of all four classes of simple policy rules that we have evaluated. Again, the numerical values reported in different cells of the table refer to the increase in the loss function—expressed in terms of inflation variability premia (IIP)—when a rule optimized in model X is used in model Y and evaluated relative to the same type of rule optimized in model Y. By this measure, two-parameter rules that respond to inflation and the output gap are the rules that are most robust to model uncertainty. The robustness properties of rules with interest rate smoothing that respond to inflation and output growth deviations from trend are slightly worse. However, this ordering can be reversed if the four-parameter rule optimized in the correct model is used as benchmark (IIP values in parentheses) and the ACEL model is dropped from the comparison. More important, model averaging helps to identify rules with interest rate smoothing and a response to output gaps that are fairly robust to model uncertainty, while regaining much of the improvement in stabilization performance promised by such rules in the absence of model uncertainty.

TABLE 5.—ROBUSTNESS OF MODEL-AVERAGING POLICY RULES
INCREASE IN IIP WHEN A MODEL-AVERAGING RULE X IS USED IN MODEL Y AND EVALUATED RELATIVE TO THE SAME TYPE OF RULE OPTIMIZED IN MODEL Y

IIP if evaluated in	2 Parameters	3 Parameters (Gap)	3 Parameters (Growth)	4 Parameters (Gaps)
	Loss(λ_y=1): $\mathrm{Var}(\pi)+\mathrm{Var}(y)+\mathrm{Var}(\Delta i)$			
SW	0.11 (1.50)[a]	1.02	0.13 (0.84)[a]	0.47
Taylor	0.03 (2.18)[a]	0.56	0.19 (1.71)[a]	1.28
ACEL	0.00 (0.17)[a]	0.27	0.40 (0.44)[a]	0.12

The values in this table concern the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a rule of the same class (2-,3-, 4-parameters) optimized in model Y. The increase is measured in terms of the implied inflation (variability) premia proposed by Kuester and Wieland (2010). The IIP corresponds to the increase in the standard deviation of the inflation rate (in percentage point terms) that would imply an equivalent increase in absolute loss.

[a]The values in parentheses refer to the increase in absolute loss in model Y under a rule optimized by averaging over all models relative to a four-parameter rule optimized in model Y.

We note that model averaging mirrors Bayesian decision making with equal prior beliefs. Kuester and Wieland (2010) compare Bayesian decision making with worst-case analysis and ambiguity aversion, which combines both objectives, in an application that deals with monetary policy modeling in the euro area. They also explore the impact of learning on posteriors and Bayesian objectives over time.

VIII. Conclusions and Extensions

The preceding comparison of the Taylor (1993a) model with the two well-known examples from the current generation of new Keynesian models of the U.S. economy by Christiano et al. (2005) and Smets and Wouters (2007) indicates a surprisingly similar monetary transmission mechanism. The empirical, model-based assessment of the impact of an unanticipated change in the federal funds rate on real U.S. GDP has not changed in the fourteen years that lie between the publication of these models. This finding is encouraging for policymakers who want to rely on such models. It differs from earlier comparison projects showing that models built later in the 1990s, such as the FRB/US model, suggested that the impact of policy shocks on real output was much more drawn out over time. Conventional wisdom on the lags of monetary policy decisions may therefore need to be revised.

The robustness analysis of simple policy rules with the three models reveals more diversity than the comparative assessment of the transmission mechanism. If the central bank has the task of stabilizing both output and inflation, then an optimal rule derived in one of our models is not robust in the other models. By sacrificing optimality in each model, one can identify some policy rules that are fairly robust, in particular, two-parameter rules that respond to

inflation and the output gap and three-parameter rules that include interest rate smoothing but replace the output gap with the deviation of GDP growth from trend.

We also find that model averaging substantially improves the robustness properties of policy rules. Hence, using a model database, such as the one described in this paper, one can derive policy rules that are more robust to model uncertainty than those obtained with a single preferred model.

Our findings also suggest at least two important extensions focusing on the implications of utility-based loss functions and a wider range of macroeconomic models.

A. Utility-Based Loss Functions

We selected the loss function in equation (6) because it has been used extensively in the past and because it corresponds to the unconditional expectation of a second-order approximation of household utility in a small New Keynesian model derived from microeconomic foundations. However, if the loss function is interpreted as a measure of utility, then its parameters (λ_y,$\lambda_{\Delta i}$) are model dependent (as we noted previously), and the list of variables appearing in the loss function must be expanded. For example, if wage rigidities are present in addition to price rigidities, not only price but also wage fluctuations will affect household utility. Onatski and Williams (2004) derive the following quadratic approximation of the unconditional expectation of household utility in the model of Smets and Wouters (2003):

$$L_{OW2004} = E\big[\pi_t^2 + 0.21K_{t-1}^2 - 0.51\pi_t\pi_{t-1} + 0.24(w_t + \pi_t)(w_t - w_{t-1})\big]. \quad (9)$$

Here w_t refers to the real wage and K_{t-1} to the lagged capital stock. To illustrate how such a loss function would affect our results, we optimized the four types of simple policy rules with respect to this utility-based loss in the SW model augmented with the variance of the change of the interest rate. Interestingly, the optimized two-, three-, and four-parameter rules have fairly similar welfare implications under the Onatski-Williams approximation of household utility in the Smets-Wouters model with a maximum difference of 1.17 in IIP terms. We also evaluate the robustness of rules optimized with respect to the simpler loss function defined by equation (6) under this new loss function. Again, model-specific two-parameter rules with inflation and the output gap, and three-parameter rules with interest rate smoothing, inflation, and output growth deviations from trend remain fairly robust, but not the other model-specific rules. More details about these results are available in the Web appendix.

B. Robustness to Other Macroeconomic Models

While we have focused on three models of the U.S. economy, the new monetary model database offers the possibility of comparing many other empirically estimated models. With regard to future research, it would be of great interest to investigate the robustness of monetary policy rules in models that offer a more detailed treatment of the financial sector. As an illustration, we extended our model comparison and robustness analysis to include the model of De Graeve (2008). De Graeve introduces a financial intermediary, capital goods producers and entrepreneurs, as in Bernanke et al. (1999), in a medium-size DSGE model of the same type as the CEE and SW models we have considered. His model, which he estimates with Bayesian methods, generates an endogenous external finance premium that is affected by a variety of economic shocks. Interestingly, we find that the GDP response to a monetary policy shock in the De Graeve (DG) model remains very close to the impulse responses in the Taylor, SW, and CEE/ACEL models reported in figure 1. The robustness of optimized model-specific rules, however, deteriorates further once we include the DG model. Especially two-parameter rules optimized in the DG model perform badly in the Taylor and SW models. However, model-averaging rules remain very robust. In fact, they need not be changed. Including the DG model in the model-averaging loss function defined by equation (8) has only a marginal effect on the optimal response coefficients in the policy rules. More information about these results is available in the Web appendix.

REFERENCES

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani, "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through," *Journal of International Economics* 72:2 (2007), 481–511.

Altig, David, Lawrence Christiano, Martin Eichenbaum, and Jesper Linde, "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," Federal Reserve Bank of Cleveland working paper no. 04-16 (2004).

Basu, Susanto, and John G. Fernald, "What Do We Know and Not Know about Potential Output?" Federal Reserve Bank of San Francisco working paper no. 09-05 (2009).

Beck, Günter, and Volker Wieland, "Central Bank Misperceptions and the Role of Money in Interest Rate Rules," *Journal of Monetary Economics* 55 (suppl. 1) (2008), S1–S17.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist, "The Financial Accelerator in a Quantitative Business Cycles Framework," in John B. Taylor and Michael Woodford (Eds.), *The Handbook of Macroeconomics*, Vol. 1C (Amsterdam: North-Holland, 1999).

Brayton, Flint, and Thomas Laubach, "Documentation of Linearized FRB/US," Board of Governors of the Federal Reserve System, Division of Research and Statistics, unpublished manuscript (2008).

Brock, William, Steven Durlauf, and Kenneth West, "Policy Evaluation in Uncertain Environments," *Brookings Papers on Economic Activity* 1 (2003), 235–302.

Bryant, Ralph, David Currie, Jacob Frenkel, Paul Masson, and Richard Portes (Eds.), *Macroeconomic Policies in an Interdependent World* (Washington, D.C.: Brookings Institution, 1989).

Bryant, Ralph, Peter Hooper, and Catherine Mann (Eds.), *Evaluating Policy Regimes: New Research in Empirical Macroeconomics* (Washington, D.C.: Brookings Institution, 1993).

Burriel, Pablo, Jesús Fernández-Villaverde, and Juan F. Rubio-Ramírez, "MEDEA: A DSGE Model for the Spanish Economy," FEDEA working paper no. 2009-17 (2009).

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

Carabenciov, Ioan, Igor Ermolaev, Charles Freedman, Michel Juillard, Ondra Kamenik, Dmitry Korshunov, and Douglas Laxton, "A Small Quarterly Projection Model of the US Economy," IMF working paper no. 08/278 (2008).

Christensen, Ian, and Ali Dib, "The Financial Accelerator in an Estimated New Keynesian Model," *Review of Economic Dynamics* 11:1 (2008), 155–178.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy* 113:1 (2005), 1–45.

Clarida, Richard, Jordi Galí, and Mark Gertler, "The Science of Monetary Policy: A New-Keynesian Perspective," *Journal of Economic Literature* 37:4 (1999), 1661–1707.

——— "A Simple Framework for International Monetary Policy Analysis," *Journal of Monetary Economics* 49:5 (2002), 879–904.

Coenen, Günter, Peter McAdam, and Roland Straub, "Tax Reform and Labour-Market Performance in the Euro Area: A Simulation-Based Analysis Using the New Area-Wide Model," *Journal of Economic Dynamics and Control* 32:8 (2008), 2543–2583.

Coenen, Günter, and Volker Wieland, "Inflation Dynamics and International Linkages: A Model of the United States, the Euro Area and Japan," ECB working paper no. 181 (2002).

——— "A Small Estimated Euro Area Model with Rational Expectations and Nominal Rigidities," *European Economic Review* 49:5 (2005), 1081–1104.

Collard, Fabrice, and Michael Juillard, "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control* 25:6 (2001), 979–1000.

De Graeve, Ferre, "The External Finance Premium and the Macroeconomy: US Post-WWII Evidence," *Journal of Economic Dynamics and Control* 32:11 (2008), 3415–3440.

Dieppe, Alistair, Keith Kuester, and Peter McAdam, "Optimal Monetary Policy Rules for the Euro Area: An Analysis Using the Area Wide Model," *Journal of Common Market Studies* 43:3 (2005), 507–537.

Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte, "Documentation of the Research and Statistics Division's Estimated DSGE Model of the U.S. Economy: 2006 Version," Board of Governors of the Federal Reserve System Finance and Economics Discussion Series working paper no. 07-53 (2007).

Erceg, Christopher J., Luca Guerrieri, and Christopher Gust, "Trade Adjustment and the Composition of Trade," *Journal of Economic Dynamics and Control* 32:8 (2008), 2622–2650.

Erceg, Christopher J., and Andrew Levin, "Optimal Monetary Policy with Durable Consumption Goods," *Journal of Monetary Economics* 53:7 (2006), 1341–1359.

Fuhrer, Jeffrey, and George Moore, "Inflation Persistence," *Quarterly Journal of Economics* 110:1 (1995), 127–159.

Galí, Jordi, and Tommaso Monacelli, "Monetary Policy and Exchange Rate Volatility in a Small Open Economy," *Review of Economic Studies* 72:252 (2005), 707–734.

Geweke, John, "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," *Econometric Reviews* 18 (1999), 1–126.

Gouvea, Solange, André Minella, Rafael Santos, and Nelson Souza-Sobrinho, "Samba: Stochastic Analytical Model with a Bayesian Approach," working paper (2008).

Gupta, Abhishek, "A Forecasting Metric for Evaluating DSGE Models," mimeograph, Johns Hopkins University (2009).

Iacoviello, Matteo, "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle," *American Economic Review* 95 (2005), 739–764.

Ireland, Peter N., "Money's Role in the Monetary Business Cycle," *Journal of Money, Credit and Banking* 36 (2004), 969–983.

Juillard, Michel, "Dynare: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables through the Use of a Relaxation Algorithm," CEPREMAP working paper no. 9602 (1996).

Klein, Lawrence, *Comparative Performance of U.S. Econometric Models* (New York: Oxford University Press, 1991).

Kuester, Keith, and Volker Wieland, "Insurance Policies for Monetary Policy in the Euro Area," *Journal of the European Economic Association* 8:4 (2010), 872–912.

Laxton, Douglas, and Paolo Pesenti, "Monetary Rules for Small, Open, Emerging Economies," *Journal of Monetary Economics* 50 (2003), 1109–1146.

Levin, Andrew, Volker Wieland, and John C. Williams, "Robustness of Simple Monetary Policy Rules under Model Uncertainty," in John B. Taylor (Ed.), *Monetary Policy Rules* (Chicago: University of Chicago Press, 1999).

——— "The Performance of Forecast-Based Monetary Policy Rules under Model Uncertainty," *American Economic Review* 93:3 (2003), 622–645.

Lucas, Robert, "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1 (1976), 19–46.

Mankiw, Gregory, and Ricardo Reis, "Sticky Information in General Equilibrium," *Journal of the European Economic Association* 5:2–3 (2007), 603–613.

McCallum, Bennett, and Edward Nelson, "Performance of Operational Policy Rules in an Estimated Semi-Classical Structural Model," in John B. Taylor (Ed.), *Monetary Policy Rules* (Chicago: University of Chicago Press, 1999).

Medina, Juan P., and Claudio Soto, "The Chilean Business Cycles through the Lens of a Stochastic General Equilibrium Model," Central Bank of Chile working paper no. 457 (2007).

Murchison, Stephen, and Andrew Rennison, "ToTEM: The Bank of Canada's New Quarterly Projection Model," Bank of Canada technical report no. 97 (2006).

Onatski, Alexei, and Noah M. Williams, "Empirical and Policy Performance of a Forward-Looking Monetary Model," University of Wisconsin, working paper (2004).

Orphanides, Athanasios, "Historical Monetary Policy Analysis and the Taylor Rule," *Journal of Monetary Economics* 50:5 (2003a), 983–1022.

——— "The Quest for Prosperity without Inflation," *Journal of Monetary Economics* 50 (2003b), 633–663.

Orphanides, Athanasios, and Volker Wieland, "Price Stability and Monetary Policy Effectiveness when Nominal Interest Rates are Bounded at Zero," Board of Governors of the Federal Reserve System Finance and Economics discussion paper no. 98-35 (1998).

Palmqvist, Stefan, "Flexible Inflation Targeting: How Should Central Banks Take the Real Economy in Consideration?" *Sveriges Riksbank Economic Review* 2 (2007).

Ratto, Marco, Werner Roeger, and Jan in 't Veld, "QUEST III: An Estimated Open-Economy DSGE Model of the Euro Area with Fiscal and Monetary Policy," *Economic Modelling* 26:1 (2009), 222–233.

Reifschneider, David, Robert Tetlow, and John C. Williams, "Aggregate Disturbances, Monetary Policy, and the Macroeconomy: The FRB/US Perspective," *Federal Reserve Bulletin,* 85:1 (1999), 1–19.

Rotemberg, Julio J., and Michael Woodford, "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual* 12 (1997), 297–346.

——— "Interest-Rate Rules in an Estimated Sticky Price Model," in John B. Taylor (Ed.), *Monetary Policy Rules* (Chicago: University of Chicago Press, 1999).

Rudebusch, Glenn D., and Lars E. O. Svensson, "Policy Rules for Inflation Targeting," in J. B. Taylor (Ed.), *Monetary Policy Rules* (Chicago: University of Chicago Press, 1999).

Schmitt-Grohe, Stephanie, and Martin Uribe, "Optimal Operational Monetary Policy in the Christiano-Eichenbaum-Evans Model of the U.S. Business Cycle," NBER working paper no. 10724 (2004).

Schorfheide, Frank, "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics* 15:6 (2000), 645–670.

Shiller, Robert J., "Comment," in L. Klein (Ed.), *Comparative Performance of U.S. Econometric Models* (New York: Oxford University Press, 1991).

Smets, Frank, and Rafael Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1:5 (2003), 1123–1175.

——— "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach," *American Economic Review* 97:3 (2007), 506–606.

Taylor, John B., "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy* 88 (1980), 1–24.

——— *Macroeconomic Policy in a World Economy* (New York: Norton, 1993a).
——— "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (1993b): 195–214.
——— (ed.), *Monetary Policy Rules* (Chicago: University of Chicago Press, 1999).
Wieland, Volker, Tobias Cwik, Gernot Mueller, Sebastian Schmidt, and Maik Wolters, "A New Comparative Approach to Macroeconomic Modelling and Policy Analysis," manuscript, Center for Financial Studies Frankfurt (2009).
Woodford, Michael, "Optimal Monetary Policy Inertia," Manchester School, supplement 67 (1999), 1–35.

APPENDIX

38 Models Included in the Model Base as of January 2011[19]

1. Small Calibrated Models	
1.1 Rotemberg and Woodford (1997)	NK_RW97
1.2 Levin, Wieland, and Williams (2003)	NK_LWW03
1.3 Clarida, Gali, and Gertler (1999)	NK_CGG99
1.4 Clarida, Gali, and Gertler 2-Country (2002)	NK_CGG02
1.5 McCallum and Nelson (1999)	NK_MCN99
1.6 Ireland (2004)	NK_IR04
1.7 Bernanke, Gertler, and Gilchrist (1999)	NK_BGG99
1.8 Galí and Monacelli (2005)	NK_GM05
2. Estimated U.S. Models	
2.1 Fuhrer and Moore (1995)	US_FM95
2.2 Orphanides and Wieland (1998)	US_OW98
2.3 FRB-US model linearized as in Levin, Wieland, and Williams (2003)	US_FRB03
2.4 FRB-US model 08 linearized by Brayton and Laubach (2008)	US_FRB08
2.5 FRB-US model 08 mixed expectations	US_FRB08mx
2.6 Smets and Wouters (2007)	US_SW07
2.7 CEE/ACEL Altig, Christiano, Eichenbaum, and Linde (2004)	US_ACELm
(m = monetary policy shock, t = technology shock, sw = SW assumptions =	US_ACELt
no cost channel, no timing constraints)	US_ACELswm
	US_ACELswt
2.8 New Fed US Model by Edge, Kiley, and Laforte (2007)	US_NFED08
2.9 Rudebusch and Svensson (1999)	US_RS99
2.10 Orphanides (2003b)	US_OR03
2.11 IMF projection model by Carabenciov et al. (2008)	US_PM08
2.12 IMF projection model with financial linkages	US_PM08fl
2.13 De Graeve (2008)	US_DG08
2.14 Christensen and Dib (2008)	US_CD08
2.15 Iacoviello (2005)	US_IAC05
2.16 Mankiw and Reis (2007)	US_MR07
3. Estimated Euro Area Models	
3.1 Coenen and Wieland (2005) (ta: Taylor-staggered contracts)	EA_CW05ta
3.2 Coenen and Wieland (2005) (fm: Fuhrer-Moore staggered contracts)	EA_CW05fm
3.3 ECB Area Wide model linearized as in Dieppe, Kuester, and McAdam (2005)	EA_AWM05
3.4 Smets and Wouters (2003)	EA_SW03
3.5. Euro Area Model of Sveriges Riksbank (Adolfson et al. 2007)	EA_SR07
3.6. Euro Area Model of the DG-ECFIN EU (Ratto, Roeger, and Veld, 2009)	EA_QUEST3
4. Estimated Small Open Economy Models (Other Countries)	
4.1 Model of the Chilean economy by Medina and Soto (2007)	CL_MS07
4.2 ToTEM model of Canada, based on Murchison and Rennison (2006), 2010 vintage	CA-ToTEM10
4.3 Model of the Brazilian economy by Gouvea et al. (2008)	BRA_SAMBA08
5. Estimated and Calibrated Multicountry Models	
5.1 Taylor (1993a) model of G7 economies	G7_TAY93
5.2 Coenen and Wieland (2002) G3 economies	G3_CW03
5.3 IMF model of euro area & CZrep by Laxton and Pesenti (2003)	EACZ_GEM03
5.4 FRB-SIGMA model by Erceg, Guerrieri, & Gust (2008)	G2_SIGMA08
5.5. ECB New-Area Wide Model of Coenen, McAdam, and Straub (2008)	EA_NAWM08

[19] See Wieland et al. (2009) for a detailed exposition of the platform for model comparison. Software and models are available for download from http://www.macromodelbase.com.