ARTURO ORMEÑO
KRISZTINA MOLNÁR

Using Survey Data of Inflation Expectations in the Estimation of Learning and Rational Expectations Models

Does survey data contain useful information for estimating macroeconomic models? We address this question by using survey data of inflation expectations to estimate the New Keynesian model by Smets and Wouters (2007) and compare its performance under rational expectations and adaptive learning. The survey information serves as an additional moment restriction and helps us to determine the learning agents' forecasting model for inflation. Adaptive learning fares similarly to rational expectations in fitting macro data, but clearly outperforms rational expectations in fitting macro and survey data simultaneously. In other words, survey data contain additional information that is not present in the macro data alone.

JEL codes: C611, D84, E30, E52
Keywords: survey data, learning, rational expectations, inflation expectations, Bayesian econometrics.

SURVEY MEASURES OF INFLATION expectations are closely followed by most central banks and are often used to examine expectations formation. (see Mankiw, Reis, and Wolfers 2004, Coibion and Gorodnichenko 2010). Even so, we know very little about how macro models fit survey data. Del Negro and Eusepi (2011) show that a dynamic general equilibrium (DSGE) model equipped with several frictions and exogenous shocks solved under rational expectations (RE), although it fits macro data well, is misspecified in fitting the dynamics of survey expectations.

An earlier version of this paper was part of Arturo Ormeño's Ph.D. dissertation at Universitat Pompeu Fabra. We would like to give special thanks to his supervisor Fabio Canova and his cosupervisor Kristoffer Nimark. We are also grateful for comments and suggestions from Marco Del Negro, Jose Dorich, Christina Felfe, Jordi Gali, Albert Marcet, Sergey Slobodyan, Robert Zymek, the editor Kenneth D. West, and two anonymous referees.

ARTURO ORMEÑO Credit Suisse AG, Switzerland; (*E-mail:* arturoormeno@gmail.com). KRISZTINA MOLNÁR Norwegian School of Economics (NHH); (*E-mail:* krisztina.molnar@nhh.no).

Received September 8, 2010; and accepted in revised form April 29, 2014.

Journal of Money, Credit and Banking, Vol. 47, No. 4 (June 2015)

This paper examines whether a DSGE model with adaptive learning can fit macro and survey data simultaneously. We use inflation survey data along with other macro variables to estimate the New Keynesian model of Smets and Wouters (2007) (SW), which is widely used by central banks for monetary policy analysis. We conduct Bayesian model comparison to contrast RE and constant-gain adaptive learning: agents form expectations about forward-looking variables by estimating linear models, using a constant-gain learning algorithm to update the parameters as new data become available. We aim to understand which expectation formation fits actual data best.

Using survey data in the estimation is a more salient issue under adaptive learning than under RE, because learning can be modeled in many ways. In an RE model, such as Del Negro and Eusepi (2011), the expectations are a function of the model parameters, and using inflation survey data can only affect the identification of the parameters. In an adaptive learning model, the dynamics of endogenous variables depend on how learning is modeled. Several papers document that adaptive learning can enhance the propagation mechanism of the DSGE models and generate the persistence that is otherwise caused by either nominal frictions or by the dynamics of the exogenous stochastic processes. However, adaptive learning can be criticized because the importance of different frictions changes depending on how learning is modeled. Milani (2007, 2011) and Slobodyan and Wouters (2012) all model learning in a different way and find different frictions to be important. To respond to this criticism, we use survey data to impose an additional consistency requirement: the learning algorithm should be chosen consistently with the observed survey data.

Our main result is that in estimations without survey data, the adaptive learning and the RE model perform similarly in fitting macro data, but when we require the model to fit survey data as well, the learning model clearly outperforms the RE model. In other words, there is information in the survey data that helps to differentiate between the different models of expectation formation. The reason behind our main result is that our benchmark specification for learning is a small (underparametrized) forecasting model: agents condition their expectations on a reduced information set compared to RE. This small forecasting model provides more time variability of inflation expectations and therefore more flexibility than RE and fits survey expectations better. This flexibility becomes apparent when we examine how the model-implied inflation expectations change when we include survey data among the observables. Expectations change substantially under learning, whereas in the RE model the model-implied inflation expectations are hardly affected. As a result, in estimations with survey data, learning captures the time pattern of the survey better than RE. We further find that big advantage of using survey data in the estimations is that it helps to avoid unrealistic jumps in model-implied inflation expectations.

We also show that the superior performance of the learning model is robust in making changes to the small forecasting model of inflation expectations. However, it is important that the learning model should be small. When we increase the model size, assuming that the agents' information set is consistent with the RE equilibrium, the dynamics of the learning model stay close to RE and the log-marginal likelihood decreases. A small forecasting model, on the other hand, is misspecified compared to RE, as a consequence the implied model dynamics are different from those of the RE model.

Besides having a small forecasting model, the time variability of expectations is also important. The likelihood decreases when we reestimate the model using a learning algorithm with less time variability (decreasing-gain learning). In this case, initial conditions have a long lasting impact on expectations; therefore, learning cannot capture episodes with large changes.

Some researchers have already pointed out the advantage of modeling learning agents with small forecasting models, because it either improves DSGE models' ability to fit macro data (Orphanides and Williams 2005, Slobodyan and Wouters 2012) or approximates inflation survey data well (Branch and Evans 2006). We show that modeling inflation expectations with a small learning model in a DSGE framework can deliver both at the same time: provide a good fit to macro data as well as observed inflation expectations.

There are only a handful of researchers who estimate macro models using data on expectations as an observable. Closest to our paper is the study by Del Negro and Eusepi (2011), which uses survey data in estimating a DSGE model exclusively under RE. It shows that survey data contain useful information to differentiate between different informational assumptions about the inflation target.[1] Milani (2011) uses survey data to show that expectation shocks, derived from the difference between survey data and an adaptive learning algorithm, can explain a major part of business cycle fluctuations. Finally, Carboni and Ellison (2009) use internal unemployment rate forecasts of the Federal Reserve in estimating the model of Sargent, Williams, and Zha (2004). Even though their focus is different from that of our paper, they also find that using data of expectations removes volatile and unrealistic beliefs implied by the model.

The remainder of the paper is organized as follows. In Section 1, we summarize the main features of the model and characterize the model's solution under both RE and learning. Section 2 presents estimation details. In Section 3, we select the benchmark learning model for inflation expectations, and describe our main estimation results and the posterior together with impulse responses. Section 4 contains robustness exercises and Section 5 concludes.

1. THE MODEL

Our estimation is based on a medium-scale New Keynesian model with price and wage rigidities, capital accumulation, investment adjustment costs, variable capital utilization, and habit formation. The model we estimate is identical to that of SW; the main innovation of our paper is to extend the SW data set by survey expectations and examine the private sector's expectation formation.

1. Del Negro and Eusepi (2011) compare full information and imperfect information RE models with fixed and time-varying inflation target.

1.1 Model Participants, Main Frictions, and Forward Variables

In this section, we give a short overview of the model without stating the equations; the detailed version can be found in the appendix in the online Supporting Information. Readers interested in more details of the model are encouraged to refer to SW.

The New Keynesian model of SW is based on a neoclassical growth model augmented with several frictions. The model consists of utility-maximizing households, profit-maximizing firms (intermediate and final goods producers), a labor union setting wages and a monetary authority.

Households. The economy is populated with a continuum of households with identical preferences that depend on hours worked and consumption. There is an external habit formation: consumption enters relative to aggregate lagged consumption. Households can move resources between periods by purchasing one-period bonds and renting capital to firms. The Euler equation describes the dynamics of consumption: consumption depends on past consumption because of habit formation, on expected future consumption because consumers prefer to smooth consumption, on expected growth in hours worked because of nonseparable preferences, and on the *ex ante* real interest rate of bonds that reflects intertemporal substitution of consumption. Households make a capital accumulation decision and decide how many units of capital services to rent to firms. Capital adjustment is costly; it is a function of the change in investment. The households' optimal investment choice is a function of past and expected future investment and the real value of the existing capital stock. The arbitrage condition for the value of the capital stock implies that this stock reacts positively to both its expected future value and the expected future real rental rate of capital, and negatively to the *ex ante* real interest rate.

Firms. There is a continuum of firms that combine capital and labor to produce differentiated intermediate goods. The model exhibits price rigidity à la Calvo (1983); therefore, firms choose their price knowing it will impact future profits. When firms cannot reoptimize their prices, they index them to past inflation. The Phillips curve of the economy is: aggregate price is a function of current and expected future marginal costs and because of indexation it is also determined by the past inflation rate. The marginal cost, in turn, is a function of the wage and the rental rate of capital.

Labor union. In order to introduce wage rigidity, SW introduce an intermediate labor union that "breaks up" households' homogenous labor supply into monopolistically differentiated labor services. The union then sets wages to maximize net present value of wages, taking into account the wage rigidity (Calvo-type) and the households' optimal decisions. After wages are set, a separate sector called labor packers buys labor from the unions, repackages it into a homogenous labor service, and resells it to the intermediate goods producers. Wages that are not reoptimized in a given period are indexed to past inflation. Aggregate wages depend on past inflation because of indexation, on expected future nominal wages because of wage rigidity, on past wages because of habit formation, and on a wage markup, which is a consequence of differentiated labor.

Monetary policy. The central bank is modeled with a Taylor rule: the interest rate is adjusted in response to inflation and to changes in the level of output. This Taylor rule is slightly different from SW, where it is a function of the output gap;[2] we decided to follow the recommendations of Slobodyan and Wouters (2009) and use output growth instead. This significantly reduces the number of forward variables, because one does not have to estimate a parallel economy under flexible prices. Slobodyan and Wouters find that this modification of the Taylor rule is innocuous in the sense that it does not change the original parameter estimates of SW.

To sum up, the model contains 13 endogenous variables: output, consumption, investment, value of the capital stock, installed stock of capital, stock of capital, inflation, capital utilization rate, real rental rate on capital, real marginal cost, real wages, hours worked, and interest rate. In addition, seven exogenous autoregressive processes are introduced, with each including an i.i.d. normally distributed error: (i) total factor productivity (in the production function), (ii) investment-specific technology (in the investment equation), (iii) risk premium (represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households), (iv) exogenous spending (in the aggregate resource constraint), (v) price markup (in the Phillips curve), (vi) wage markup (in the aggregate wage equation), and (vii) a monetary policy shock (in the Taylor rule).

The model is detrended with respect to the deterministic growth rate of the labor-augmenting technological progress and linearized around the steady state of the detrended variables. The set of equations that describe the dynamics of this model can be assembled into the following two equations:

$$\Theta_0 \widetilde{E}_t Y_{t+1} + \Theta_1 Y_t + \Theta_2 Y_{t-1} + \Psi e_t = 0 \tag{1}$$

$$e_t = \Gamma_e e_{t-1} + \Gamma_\varepsilon \varepsilon_t, \tag{2}$$

where Y is a vector containing the 13 endogenous variables of the model, e is the vector of the seven exogenous shocks, and ε is the vector of i.i.d. normal innovations. $\widetilde{E}_t(\cdot)$ denotes expectations that are not necessarily rational. The matrices Θ_0, Θ_1, Θ_2, and Ψ contain the nonlinear combinations of the model parameters. Zero elements of Θ_0 and Θ_2 correspond to variables that are not present in the model with lagged or expected future values (see the appendix in the online Supporting Information). Γ_e is a diagonal matrix that contains the autoregressive coefficients of the exogenous shocks. Since almost all innovations are independent, Γ_ε is an identity matrix except for one off-diagonal element: in the estimations, we assume that productivity innovation can affect the spending shocks.[3]

2. That is, the difference between the output obtained under nominal rigidities and under flexible prices.

3. We follow SW to introduce this off-diagonal element because exogenous spending data include net exports, which may be affected by domestic productivity developments.

1.2 The Rational Expectation Solution

When we estimate the model under RE, we assume that private agents have perfect knowledge about the model, its parameters, and the true stochastic processes of the economy. We use Uhlig (1999) to solve (1) and (2).[4] We focus on determinate RE solutions, and restrict the parameter space accordingly. The resulting law of motion takes the following form:

$$Y_t = \Phi_1^{RE} Y_{t-1} + \Phi_2^{RE} e_{t-1} + \Phi_3^{RE} \varepsilon_t. \tag{3}$$

1.3 The Adaptive Learning Solution

The second expectation formation we examine is adaptive learning in the sense of Bray (1982), Marcet and Sargent (1989), and Evans and Honkapohja (2001). Adaptive learning is popular in estimations of structural models because it often improves the model fit of standard RE models (see, e.g., Slobodyan and Wouters 2009, Milani 2007). This departure from RE is often motivated by the argument that the level of cognitive ability and computational skill required by the RE assumption is too high.

Adaptive learning agents behave like econometricians and use estimations to forecast future variables. We assume that agents estimate a linear function consistent with the RE solution (3). They estimate a system of linear equations $Y_t = \beta_t X_{t-1} + \varepsilon_t^y$, where β_t is a matrix of coefficient estimates and X_{t-1} includes lagged values of those endogenous variables Y and exogenous shocks e that agents use in their estimations. In practice, agents might not use all endogenous variables and exogenous shocks. The precise choice of regressors is described in Section 3.1.[5]

Agents generate forecasts in the following way. At time t they use data up to time $t-1$, and obtain coefficient estimates β_{t-1}; they then forecast the vector of endogenous variables as

$$E_t^{LS} Y_{t+1} = \beta_{t-1} X_t, \tag{4}$$

where $E_t^{LS}(\cdot)$ denotes expectations of learning agents. We assume that agents observe contemporaneous data; therefore, X_t includes contemporaneous values of the regressors.[6]

4. Alternative solution algorithms can be found in Blanchard and Kahn (1980), Binder and Pesaran (1997), Christiano (2002), and Sims (2002).

5. In the estimations we assume that agents also use a constant. In the RE solution of the model there is no constant; in the data, however, detrended variables might not have a zero average, and this can be captured by including a constant. See more about this in Section 2, which describes the estimation.

6. The same timing assumption is used in Canova and Gambetti (2010). Alternatively, one could assume that at time t agents only observe Y_{t-1}. Del Negro and Eusepi (2011) examine different timing assumptions and find that their model fit is not affected by the timing assumption. A consequence of our timing assumption is that after substituting (4) into (1), Y_t appears on both sides; therefore, in the estimations we have to rearrange these equations.

We assume that agents update their coefficient estimates with a constant gain least squares (CG-LS)

$$\beta_t = \beta_{t-1} + g R_t^{-1} X_{t-1}(Y_t - \beta_{t-1} X_{t-1})' \tag{5}$$

$$R_t = R_{t-1} + g(X_{t-1} X_{t-1}' - R_{t-1}), \tag{6}$$

where R_t is the variance covariance matrix of the stacked regressors and $g \in [0, 1)$ is the constant-gain *tracking parameter*. It can be shown that when the sample is large enough, these recursions approximate the exponentially weighted least squares estimate (see Sargent 1999). The higher the constant-gain parameter g, the more responsive is the coefficient estimate β to new data. Therefore, a high gain parameter helps to track structural changes, but at the same time, it does not filter the noise from the data as well.[7] We believe constant-gain learning is a desirable way to model expectations because it fits U.S. surveys well and at the same time it provides good forecasts of macro data (see Branch and Evans 2006).[8]

After substituting (4) into (1) and rearranging we get the equilibrium under learning

$$Y_t = \Phi_{1,t-1}^{LS} Y_{t-1} + \Phi_{2,t-1}^{LS} e_{t-1} + \Phi_{3,t-1}^{LS} \varepsilon_t. \tag{7}$$

The matrices $\Phi_{1,t-1}^{LS}$, $\Phi_{2,t-1}^{LS}$, $\Phi_{3,t-1}^{LS}$ depend on the model coefficients and the coefficient estimates β_{t-1}. Since β_{t-1} is reestimated in each time period, its presence introduces time variability into the coefficient matrices of (7).

2. DATA AND ESTIMATION DETAILS

In this section, we describe the data set (2.1), estimation details under RE and adaptive learning (2.2), and the Bayesian estimation (2.3).

2.1 Data

We estimate the SW model on U.S. data. For comparability, we use the same data set as SW, but we extend it with survey expectations of inflation.

Our quarterly macroeconomic indicators are: real GDP (*GDP*), real consumption (*Cons*), real investment (*Inv*), real wage (*Wage*), GDP deflator (*P*), hours worked (*Hours*) and the federal funds rate (*FedFunds*). All variables are expressed in log difference (*dl*) except one variable, hours worked, which is in logarithm (*l*) and the federal funds rate which is not transformed. Please refer to the online appendix for a detailed description of the data.

7. For more on optimal adaptive algorithms, see Benveniste, Métivier, and Priouret (1990).
8. Branch and Evans (2006) examine inflation and GDP surveys.

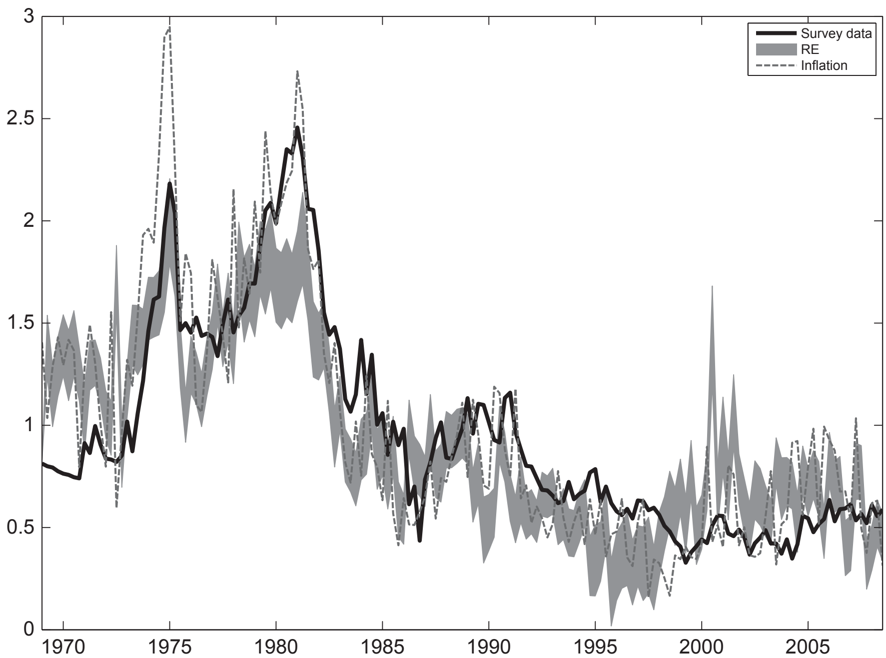


FIG. 1. Inflation, Survey Data, and Model-Implied Expectations under RE (Database without Survey Data).

NOTES: The model-implied inflation expectations are obtained using the Kalman-filtered estimates at each set of parameter values that conforms the posterior distributions. The gray and black areas represent the distance between the 5th and 95th percent confidence bands.

Our survey measure is the GDP deflator from the Survey of Professional Forecasters (SPF), which is collected quarterly. Our sample stretches from the beginning of the SPF survey 1968:4 until 2008:2.[9] Figure 1 plots the median SPF inflation survey; we denote this series by P^{SPF}. The survey series is log-differenced, similarly to the real GDP deflator data, to express one-quarter-ahead expectations. We use only the GDP deflator survey data because all other SPF data have either been collected only from a later date or are not a forward variable in our model.[10]

2.2 *Estimation Details*

In this subsection, we describe the state-space representation of the DSGE model that we estimate with the Kalman filter. Our measurement equations that relate our

9. This sample period is about the same as the SW sample, but slightly shorter.

10. Personal consumption expenditures, and residential and nonresidential fixed investment were collected only from 1981. Real GDP expectations can be constructed from 1968 but they are not a forward variable in our model. We use only SPF, because it is the longest survey. The Michigan survey is 10 years shorter than the SPF, and the Livingstone survey is only biannual.

macroeconomic data to the variables of the model under both RE and CG-LS (apart from the measurement equation of survey data) are

$$\begin{bmatrix} dlGDP_t \\ dlCons_t \\ dlInv_t \\ dlWage_t \\ lHours_t \\ dlP_t \\ FedFunds_t \end{bmatrix} = \begin{bmatrix} \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{l} \\ \bar{\pi} \\ \bar{r} \end{bmatrix} + \begin{bmatrix} \hat{y}_t - \hat{y}_{t-1} \\ \hat{c}_t - \hat{c}_{t-1} \\ \hat{i}_t - \hat{i}_{t-1} \\ \hat{w}_t - \hat{w}_{t-1} \\ \hat{l}_t \\ \hat{\pi}_t \\ \hat{R}_t \end{bmatrix}, \tag{8}$$

where $\bar{\gamma}$ is the common quarterly trend growth rate, $\bar{l}$ is steady-state hours worked, $\bar{\pi}$ is the steady-state quarterly inflation rate, and $\bar{r}$ is the steady- state quarterly nominal interest rate.

Estimations under RE. Under RE the law of motion of the exogenous shocks (2) and the solution under RE (3) form a state-space representation of the DSGE model. When we use survey data along with macro data, we supplement (8) with an additional measurement equation for expectations:

$$dlP^{SPF}_{t,t+1} = \bar{\pi} + E_t\hat{\pi}_{t+1} + \zeta_t = \bar{\pi} + \Phi^{RE}_{1,\pi}Y_t + \Phi^{RE}_{2,\pi}e_t + \zeta_t, \tag{9}$$

where $dlP^{SPF}_{t,t+1}$ is the log difference of SPF inflation expectations from quarters t to $t+1$, $\Phi^{RE}_{1,\pi}$ and $\Phi^{RE}_{2,\pi}$ are the rows of the RE model solution (3) that correspond to inflation, and ζ_t is an i.i.d. measurement error. The interpretation of (9) is that survey data are viewed as a noisy measure of model-consistent RE.

When survey data are used as an observable, it introduces cross-equation restrictions: agents' rational inflation expectations have to explain not only the model equations containing inflation expectations (Phillips curve, consumption Euler equation, wage equation, and equation for the value of the capital stock), but also the SPF survey.

We assume that agents use the Kalman filter to estimate latent variables. These include some endogenous variables, because they are not in our data set, as well as the exogenous shocks.

Estimations under learning. When we model agents with CG-LS, the state-space representation of the model consists of the law of motion of the stochastic shocks (2) and the solution under CG-LS (7).

Our forward variables are inflation, consumption, investment, hours worked, real wages, real rental rate on capital, and the value of the capital stock. To determine a CG-LS forecasting model ((5)–(6)) for each of these we respond to a general criticism of nonrational expectations, namely, the abundance of forecasting models available to chose from. We explain the details below.

We use survey data to discipline the model choice for *inflation expectations*. We search for the CG-LS model that provides the best fit to the survey data in terms of one-quarter-ahead out-of-sample forecasts. Our method is explained in detail in Section 3.1. We estimate initial conditions for the CG-LS recursion from presample data.

For *other forward variables* we do not have survey data; we therefore model CG-LS agents to be as close as possible to the RE equilibrium. We assume that agents use the correct minimum state variable (MSV) representation (3) to estimate their regressions; that is, they condition their expectations on the same information set as rational agents, but they do not know the coefficients of the RE equilibrium and therefore have to estimate it. This way, we do not make *ad hoc* assumptions about these expectations, but use the RE model solution instead.[11] In our model, this implies that agents use a state vector containing 20 variables, many of which are unobserved (latent). We follow Slobodyan and Wouters (2012) and assume that learning agents are similar to RE agents, and use the Kalman filter recursion to generate estimates of the latent variables. We set β_0 equal to the coefficients of the RE solution (3) and the initial condition of R_0 equal to the unconditional second moments resulting from the RE solution. In the estimations, we implement this by solving for the RE equilibrium for each draw of parameters during the Bayesian estimation.[12] Note that if the estimated tracking parameter is zero, the CG-LS coefficients are never updated and stay equal to their initial RE value; thus, learning forecasts are identical to the RE forecasts. If the estimated tracking parameter is nonzero, new data affect the β_t estimates (through the learning recursion (5)–(6)) and learning forecasts would deviate from the RE forecasts. This way learning introduces new long-run perpetual dynamics, different from that of RE.

Since we are modeling inflation expectations and expectations of other forward variables differently, we have two sets of the CG-LS recursions ((5)–(6)) and we estimate a different tracking parameter for each.

In the estimations using the SPF inflation survey as observable, we estimate a new measurement equation for expectations:

$$dlP^{SPF}_{t,t+1} = \bar{\pi} + E^{LS}_t \pi_{t+1} + \zeta_t, \tag{10}$$

where $E^{LS}_t \pi_{t+1}$ is the row corresponding to inflation in learning agents' expectation, equation (4).[13] The interpretation of equation (10) is that we treat SPF inflation survey as a noisy measure of CG-LS inflation expectations.

11. Strictly speaking, expectations that are rational in the RE model are not rational when inflation expectations are CG-LS. RE within the model should take into account that inflation is formed with CG-LS learning. We would like to thank an anonymous referee for pointing this out.

12. We opted not to estimate the initial conditions together with the other coefficients, because it would increase the number of parameters we have to estimate. Also, treating initial conditions as free parameters to be estimated might be problematic because initial conditions might end up explaining a too large portion of the model fit. (See Carboni and Ellison 2009 citing Sims' critique of Sargent, Williams, and Zha 2004.) For a paper estimating the initial conditions, see Sargent, Williams, and Zha (2004) and Slobodyan and Wouters (2012).

13. We implicitly assume that inflation and survey expectations have the same steady state.

When survey data are used as an observable in the estimations of the learning model, the model solution under learning (7) and the measurement equation of inflation expectations (10) both depend on the CG-LS coefficient estimates. This introduces similar cross-equation restrictions as the estimation under RE: agents' inflation expectations have to explain not only those model equations that contain inflation expectations but also the SPF survey.[14]

2.3 Parameters, Priors, and Bayesian Estimation

The structural model contains 38 parameters. We estimate 33 of these and for the remaining 5 we follow SW and use calibrated values.[15] In the learning estimation, we have two additional parameters (the gain for inflation expectations and the gain for all other expectations). In estimations using survey data the standard deviation of the measurement error of expectations ζ_t is an extra parameter.

The prior distributions of the structural parameters are the same as in SW (see Table 1). We use uniform distributions over the [0,1) domain for the gains and an inverse gamma distribution with a zero mean and a standard deviation of two for the standard deviation of ζ_t.

The DSGE model is estimated using Bayesian estimation methods. Employing the random walk Metropolis–Hastings algorithm, we obtain 250,000 draws from each model's posterior distribution. The first half of these draws is discarded, and 1 out of every 10 remaining draws is selected to estimate the moments of the posterior distributions.

3. RESULTS

In this section, we describe how we use the SPF survey to select the CG-LS model for inflation expectations (3.1), then show our main results (3.2), and discuss parameter estimates and impulse response functions (3.2.2).

3.1 Forecasting Models for Inflation

To determine the CG-LS model for inflation expectations, we search for the model that generates fictional out-of-sample inflation forecasts that are the closest to the SFP inflation survey.

Let us denote the set of possible forecasting models by Ω_X and an element, one specific model, by X^i. For each forecasting model X^i and tracking parameter g we

14. Ireland (2003) advocates deriving cross-equation restrictions under learning in the same way as under RE. Another example is Carboni and Ellison (2009) who introduce cross-equation restrictions in a model where the central bank is using a Kalman filter to learn about the Phillips curve.

15. The calibrated parameters are: depreciation rate, exogenous spending–GDP ratio, steady-state markup in the labor market, and two parameters of the function that aggregates differentiated labor and output (see the online appendix for values). The first two of these are difficult to identify in estimations, while the last three are not identified (see SW).

TABLE 1

Prior Distributions of Structural Parameters

	Symbol	Distribution	Mean	Std.
Share of capital in production	α	Normal	0.30	0.05
Inv. elasticity of intertemporal substitution	σ_c	Normal	1.50	0.38
Fix cost in production	Φ	Normal	1.25	0.13
Adjust cost of investment	S''	Normal	4.00	1.50
Habits in consumption	η	Beta	0.70	0.10
Wage stickiness	ξ_w	Beta	0.50	0.10
Inv. elast. labor supply	σ_l	Normal	2.00	0.75
Price stickiness	ξ_p	Beta	0.50	0.10
Wage indexation	ι_w	Beta	0.50	0.15
Price indexation	ι_p	Beta	0.50	0.15
Capital utilization elasticity	ψ	Beta	0.50	0.15
Taylor rule: response to inflation	r_π	Normal	1.50	0.25
Taylor rule: response to lagged interest rate	ρ_R	Beta	0.75	0.10
Taylor rule: response to changes in output	$r_{\Delta y}$	Normal	0.13	0.05
Trend growth rate	$\bar{y}$	Normal	0.40	0.10
Steady state of inflation	$\bar{\pi}$	Gamma	0.63	0.10
Steady state of hours worked	$\bar{l}$	Normal	0.00	2.00
Steady state of nominal int rate	$\bar{r}$	Gamma	1.15	0.30
Autocorrelation coef. price markup shock	ρ_p	Beta	0.50	0.20
Autocorrelation coef. wage markup shock	ρ_w	Beta	0.50	0.20
Autocorrelation coef. product. shock	ρ_a	Beta	0.50	0.20
Autocorrelation coef. risk premium shock	ρ_b	Beta	0.50	0.20
Autocorrelation coef. government shock	ρ_g	Beta	0.50	0.20
Autocorrelation coef. investment-specific shock	ρ_q	Beta	0.50	0.20
Autocorrelation coef. monet policy shock	ρ_r	Beta	0.50	0.20
Correlation government and productivity shocks	ρ_{ga}	Normal	0.50	0.25
Std price markup innovation	σ_p	Inv. Gamma	0.10	2.00
Std. wage markup innovation	σ_w	Inv. Gamma	0.10	2.00
Std. product. innovation	σ_a	Inv. Gamma	0.10	2.00
Std. risk premium innovation	σ_b	Inv. Gamma	0.10	2.00
Std. government innovation	σ_b	Inv. Gamma	0.10	2.00
Std. inv. specific innovation	σ_q	Inv. Gamma	0.10	2.00
Std. monet policy innovation	σ_r	Inv. Gamma	0.10	2.00
Gain—others than inflation	$g^{\text{non}\pi}$	Uniform	0.00	1.00
Gain—inflation	g^{π}	Uniform	0.00	1.00
Std. measurement error on expectations	σ_{exp}	Inv. Gamma	0.10	2.00

generate one-quarter-ahead CG-LS inflation forecasts in the following way. First, we use presample estimates of the model to set the initial conditions for the learning recursion. We have chosen a long presample, 1950:1–1968:3, to avoid oversensitivity of the initial presample estimates of β_0, R_0. The end point 1968:3 is given by the start of the SPF data. In the next step, from 1968:4 onward we use new data and the learning recursion ((5)–(6)) to update β and then generate one-period-ahead forecasts.[16] We

16. As previously discussed, agents use the current quarter's data to forecast. In reality, survey respondents do not know the current quarterly data but have a good estimate of it. The SPF survey is conducted in the middle of the quarter, when preliminary data are already published. Del Negro and Eusepi (2011) examine the importance of this timing assumption and find that estimation results are not sensitive to it (contemporaneous timing versus agents accessing only to $t-1$ data).

TABLE 2

RANKING OF FORECASTING MODELS FOR INFLATION

Rank	Regressors	Gain	MSD
1	*dlP*	0.125	0.0294
2	*dlP lHours*	0.113	0.0300
3	*dlP dlCons*	0.100	0.0302
4	*dlP dlCons lHours*	0.125	0.0303
5	*dlP dlGDP*	0.125	0.0315

NOTES: The models are estimated by recursive CG-LS. The initial conditions are obtained from the period 1950:1–1968:3. Regression: $dlP_t = \beta_0^\pi + \beta_1^\pi \text{regressor}_{t-1}$; Sample period 1968:4–2008:2. MSD = mean squared deviation from survey expectations.

choose the value of g and the set of regressors X^i that generate inflation forecasts closest in mean squared deviation from the survey expectations:

$$\underset{0 \leqq g < 1, X^i \in \Omega_X}{\operatorname{argmin}} \sum_{t=1}^{T} \left(\pi^e_{t,t+1}(g, X^i) - dlP^{SPF}_{t,t+1}\right)^2, \tag{11}$$

where $1 \ldots T$ is the survey sample 1968:4–2008:2 and $dlP^{SPF}_{t,t+1}$ is the survey measure of one-period-ahead inflation expectations conducted at time t.

The set of possible regressors consists of all the data we use in the estimations: *dlGDP*, *dlCons*, *dlInv*, *dlWage*, *dlP*, *FedFunds*, and *lHours*. In other words, we assume that agents have access to the same data set as we have. We also include a constant in the set of regressors to account for the fact that actual inflation is not zero on average.

In total we estimate 127 models, each with a grid of different tracking parameters in [0, 1). The regressors that minimize (11) and provide the best fit to the SPF inflation survey are lagged inflation and a constant (see Table 2). Even though this result suggests that professional forecasters use only past inflation to forecast inflation, we can also see from Table 2 that other forecasting models that include real economic variables provide a very similar fit. Figure 2 shows that the five best-fitting models are indeed hardly distinguishable, and they all track the time series of inflation survey well.

The estimated tracking parameters are high, for the best-fitting model $g = 0.12$. This suggests that agents use about 2 years of data for their inflation forecast.[17] On U.S. data, smaller gain parameters were estimated in papers that modeled learning with a vector autoregression (VAR) because VAR learning is very unstable with a high gain parameter (see Branch and Evans 2006, Milani 2007, Slobodyan and Wouters 2012). We find that the small linear forecasting models in Table 2 do not become unstable with a high g; moreover, a high-gain parameter allows for substantial time variability in the learning coefficients that describes observed inflation expectations well.

17. This is calculated with $1/g$, which is the halving time of CG-LS; that is, beyond this period less than 50% of the data are used for the estimations.

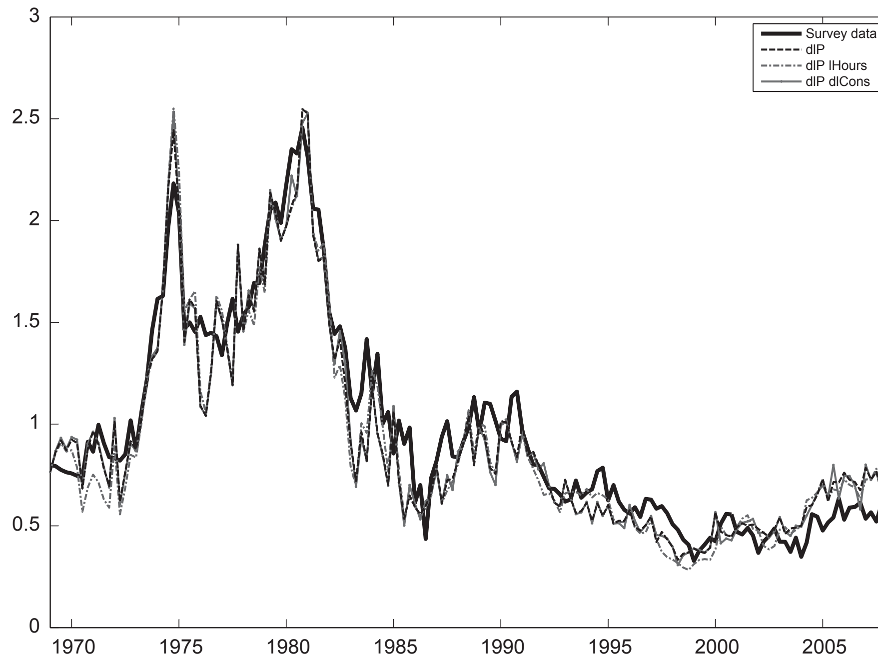


FIG. 2. Inflation Forecasts and Survey Data of Inflation Expectations.

NOTES: The model-implied inflation expectations are obtained using the Kalman-filtered estimates at each set of parameter values that conforms the posterior distributions. The gray and black areas represent the distance between the 5th and 95th percent confidence bands.

The generated CG-LS inflation forecasts fit particularly well the increase in inflation expectations in the 1970s, when monetary policy was used to raise inflation in the belief that there was an exploitable trade-off between inflation and unemployment.[18] A high tracking parameter is essential for this fit. Agents' initial belief was estimated from a low-inflation presample period and a high g implies that agents "understand" sooner that in the 1970s a new high inflation period has started; they adjust their expectations more quickly in response to changes in the data, and inflation expectations converge further away from their initial belief.

Our benchmark CG-LS model for inflation expectations is the best-fitting model

$$E_t^{LS} dlP_{t+1} = \beta_{0,t-1}^{\pi} + \beta_{1,t-1}^{\pi} dlP_t, \tag{12}$$

where $\beta_{0,t-1}^{\pi}$ and $\beta_{1,t-1}^{\pi}$ are the time t estimates of the constant and the coefficient of lagged inflation, respectively. As we assume contemporaneous timing, agents use

18. At the same time the unemployment rate increased as well, so the traditional Phillips curve inflation–output relationship broke down.See more on this and alternative explanations for the behavior of inflation in the 1970s in Cogley and Sargent (2002).

TABLE 3

MODEL COMPARISON

Log-marginal likelihood	Data set without survey data (1)	Data set with survey data (2)	(3) = (2) – (1)
RE	−146.78	−19.14	127.64
Learning	−142.82	45.22	188.04

NOTES: This table shows the log marginal likelihood of the RE and learning model. Survey data: median of SPF one-quarter-ahead forecast of the GDP deflator.

their time t estimate and the current value of inflation dlP_t to forecast inflation at $t + 1$. We use this benchmark specification (12) for our main estimations, and in the robustness section (Section 4), we examine other forecasting models in Table 2 as well.

3.2 Model Fit and Model-Implied Inflation Expectations

Our main estimation results are summarized in Table 3.

In the estimations without survey data (column (1)) the CG-LS model performs slightly better, but the difference of log marginal likelihoods is only 3.96 points.[19,20] However, when SPF inflation expectations are included among the observables, the CG-LS model clearly outperforms the RE model (Table 3, column (2)). The difference between their log-marginal likelihoods is 64.36 points, which implies a very high posterior odd of 8.93E+27. This likelihood difference in favor of the adaptive learning specification is much larger than in earlier studies that did not use survey data among the observables. For example, in Slobodyan and Wouters (2009, 2012), the marginal likelihood of learning is only 10–20 points higher than that of RE.[21] The fact that the difference in the likelihood of the RE and the learning model increases so much implies that there is extra information in the survey as to which model describes the reality best.

Learning outperforms RE due to the flexibility of the learning model of inflation expectations. To illustrate this, Figure 3 plots how much the model-implied inflation expectations change when survey data are added to the observables. Inflation expectations implied by the RE model do not change much: differences in the survey and

19. The difference is even smaller with uniform prior distributions: log marginal likelihood of RE and learning is –120 and –119.2, respectively. Del Negro and Schorfheide (2008) find that even a five-point difference in the log marginal likelihood can be overturned by choosing a slightly different prior.

20. Although previous literature found that modeling agents as learners can improve the likelihood, we are not surprised that in our estimations this is not the case. As Slobodyan and Wouters (2012) show, the likelihood depends on how learning is modeled. In general, learning with small forecasting models provides a better fit, but we use a small forecasting model only for learning about inflation.

21. Similar to us, Del Negro and Eusepi (2011) find that including survey data among the observables can increase the log marginal likelihood by a similar magnitude. Their analysis is different from ours because they do not examine adaptive learning.

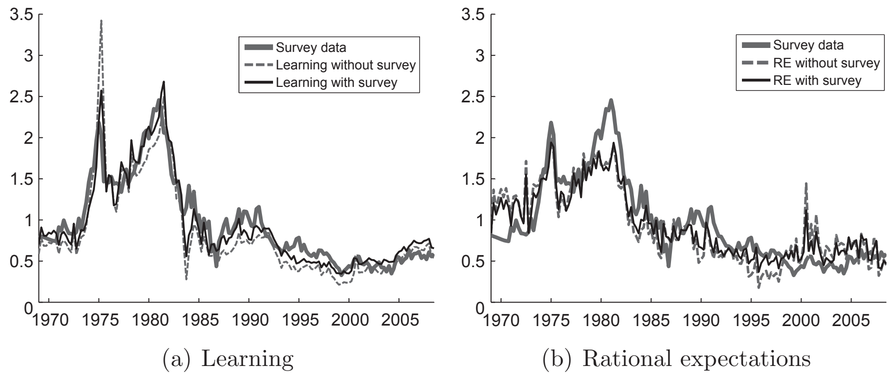


FIG. 3. Model-Implied Inflation Expectations, with or without Survey Data in the Database.

the model-implied inflation expectations end up mostly in the error term. Expectations implied by learning, on the other hand, change more and become closer to the survey data. This provides a better likelihood both because the survey of inflation expectations is explained better and because, at the same time, the learning model also fits macro variables well.

The reason behind this flexibility of adaptive learning is the small, linear forecasting model of inflation. This small forecasting model is misspecified and not consistent with the RE equilibrium, and therefore the model dynamics under learning and RE are markedly different. In addition, due to the high estimated gain parameter (see posterior estimates, Table 4), the CG-LS coefficients vary substantially in response to new data and have considerable time variability (see estimated CG-LS coefficients in Figure 4).

In order to disentangle the effect of a small forecasting model from adaptive learning we reestimate our DSGE model changing the small forecasting model of inflation expectations into a model consistent with the MSV solution under RE.[22] In our medium-scale DSGE model, this implies a large forecasting model. The first row of Table 5 shows that in estimations with survey data, the learning model still has a higher log marginal likelihood than RE but, compared to the small model in Table 3, MSV learning worsens the likelihood. MSV learning implies similar dynamics for inflation expectations as RE, but with more time variability (Figure 5). Similar to RE, MSV learning fails to capture the increase of inflation expectations in the 1970s and their decline in the 1980s. We conclude that learning can improve the model fit, but a small forecasting model is essential to attain the good performance. (In Section 4, we examine the robustness of this result to changing the small forecasting model.)

A small forecasting model in itself is not enough for a good model fit, but time variation of the learning coefficients, caused by the high estimated gain parameter, is

22. We would like to thank an anonymous referee for suggesting this estimation.

TABLE 4

POSTERIOR DISTRIBUTION STATISTICS

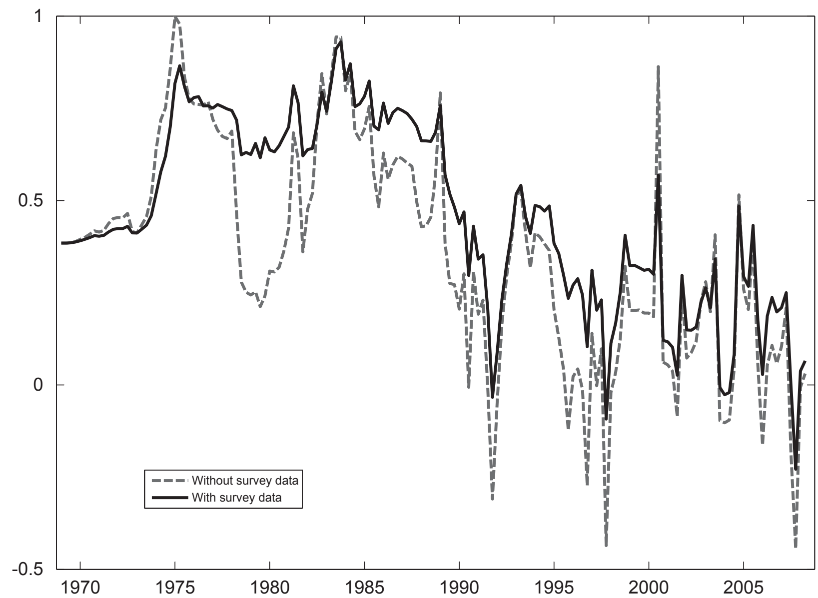
		(1)		(2)		(3)		(4)	
		Without survey data				With survey data			
		RE		Learning		RE		Learning	
Parameter	Symbol	Median	Std.	Median	Std.	Median	Std.	Median	Std.
Wage stickiness	ξ_w	0.554	0.045	0.547	0.049	0.468	0.043	0.563	0.049
Price stickiness	ξ_p	0.648	0.044	0.481	0.035	0.629	0.058	0.480	0.035
Wage indexation	ι_w	0.482	0.131	0.314	0.107	0.442	0.124	0.319	0.107
Price indexation	ι_p	0.327	0.155	0.544	0.108	0.052	0.025	0.515	0.119
TR: inflation	r_π	1.666	0.130	1.396	0.116	1.711	0.114	1.398	0.104
TR: lag interest rate	ρ_R	0.760	0.028	0.763	0.028	0.706	0.030	0.777	0.029
TR: change in output	$r_{\Delta y}$	0.199	0.046	0.203	0.046	0.187	0.044	0.210	0.047
Aut. price markup shock	ρ_p	0.448	0.195	0.140	0.070	0.726	0.078	0.173	0.087
Std. price markup shock	σ_p	0.145	0.026	0.213	0.017	0.112	0.013	0.204	0.014
Gain—inflation	g^π			0.188	0.014			0.141	0.009
Gain—others	$g^{\text{non}\pi}$			0.031	0.042			0.019	0.031
Measurement exp error	σ_{exp}					0.265	0.016	0.176	0.010
Log. mg. likelihood		−146.8		−142.8		−19.1		45.2	

NOTES: This table shows the median and standard deviation of the posterior distributions of the parameters most closely related to inflation dynamics. (TR= coefficients of the Taylor rule.) The online appendix contains the same statistics for the complete list of parameters of the model, their prior and posterior distributions, and a convergence check of the random walk Metropolis–Hastings.

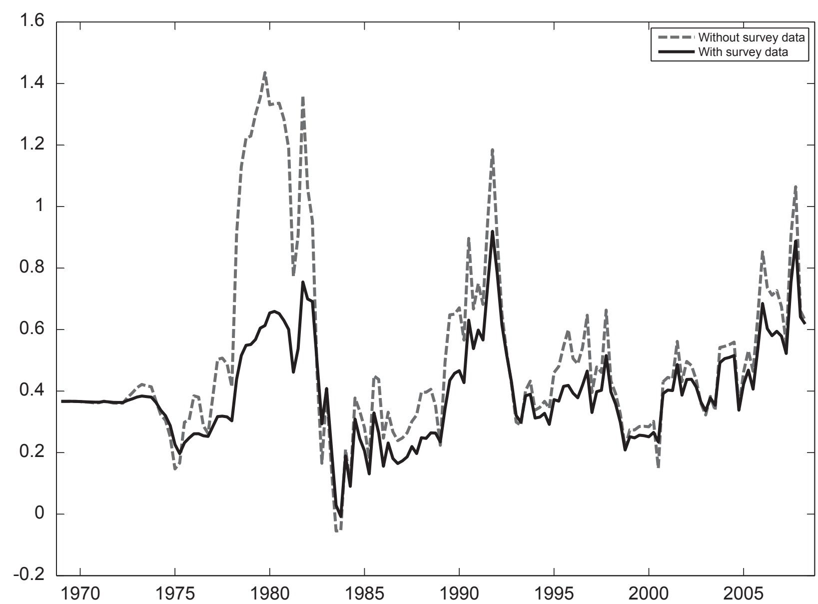
also important. To show this we reestimate the model using decreasing-gain learning (DG-LS) instead of constant-gain learning. Under DG-LS the gain parameter is decreasing in time, $g = 1/t$; thus, in time the learning recursion reacts less and less to new data. The second row of Table 5 shows that when we include survey data among the observables, the likelihood worsens so much that the RE model now outperforms DG-LS. The intuition behind is that with DG-LS it takes a long time for the initial conditions to die out (see Figure 5). Since the initial conditions are obtained during a period of low and not persistent inflation (period 1950:1–1968:3), the model fails to replicate the increases in SPF inflation expectations in the 1970s and their substantial decline.

Goodness of fit to inflation survey data. To quantify the fit to observed survey expectations, we follow the method of Del Negro and Eusepi (2011). Let $\mathcal{M}_i$ denote the model (solved either under RE or CG-LS) and $Y_{1,T}$ the set of macroeconomic observables used in the estimation. As before, dlP^{SPF} is the one-quarter-ahead SFP inflation forecast (level forecasts are log-differenced). Subindex $1, T$ denotes the data sample $t = 1 \ldots T$. We are interested in the posterior likelihood $p(dlP^{SPF}_{1,T}|Y_{1,T}, \mathcal{M}_i)$ that provides information on how well model $\mathcal{M}_i$ estimated with the macroeconomic data $Y_{1,T}$ fits the survey data. An easy way to compute this quantity is

$$p(dlP^{SPF}_{1,T}|Y_{1,T}, \mathcal{M}_i) = \frac{p(dlP^{SPF}_{1,T}, Y_{1,T}|\mathcal{M}_i)}{P(Y_{1,T}|\mathcal{M}_i)}. \tag{13}$$



(a) Perceived inflation persistence



(b) Constant

Fig. 4. Time Variation in the Learning Coefficients of Inflation Expectations.

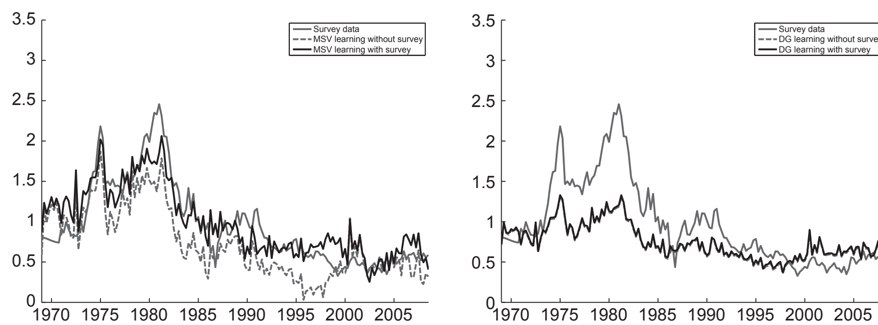
Notes: Regression: $dlP_t = \beta_0^{\pi} + \beta_1^{\pi} dlP_{t-1}$. (Best-fitting model to SPF inflation survey.)

TABLE 5

Model Comparison, Estimation Using Decreasing-Gain (DG) Learning, and MSV Learning

Log marginal likelihood	Data set without survey data (1)	Data set with survey data (2)	(3) = (2) – (1)
MSV learning	−131.9	−5.0	126.9
DG learning	−148.35	−77.20	71.14

Notes: This table shows the log marginal likelihood of the RE and learning model. Survey data: median of SPF one-quarter-ahead forecast of the GDP deflator.



(a) MSV Learning (b) DG Learning

Fig. 5. Model-Implied Inflation Expectations, with or without Survey in the Database.

The numerator and the denominator of (13) are reported in logarithm in columns (2) and (1) of Table 3. In column (3) we report the logarithm of (13), which is the difference of columns (2) and (1). Our result shows that learning clearly outperforms RE in describing the evolution of the survey data.

Learning fits the survey well on average and it provides a better fit than RE across time (see Figure 3). In comparison to RE, expectations generated by the learning model follow the survey much more closely, especially during the period of high variance in the first half of our sample. This can be explained by the fact that the learning algorithm estimates high inflation persistence in the 1970s and 1980s (see Figure 4(a)); therefore, learning expectations increase together with inflation in the 1970s and decline together with inflation in the 1980s Volker era. This pattern describes survey expectations well with one notable exception. In 1983, the learning model forecasts a decrease in inflation expectations when in fact the survey expectations increase due to the uncertainty around Volker's reappointment by Reagan (coupled with a high budget deficit). The CG-LS algorithm forecasts a decrease in inflation expectations, because actual inflation decreased during this time. In other words, the uncertainty around Volker's reappointment was an exogenous information that survey respondents used but CG-LS estimates could not, because this uncertainty was not visible in the data.

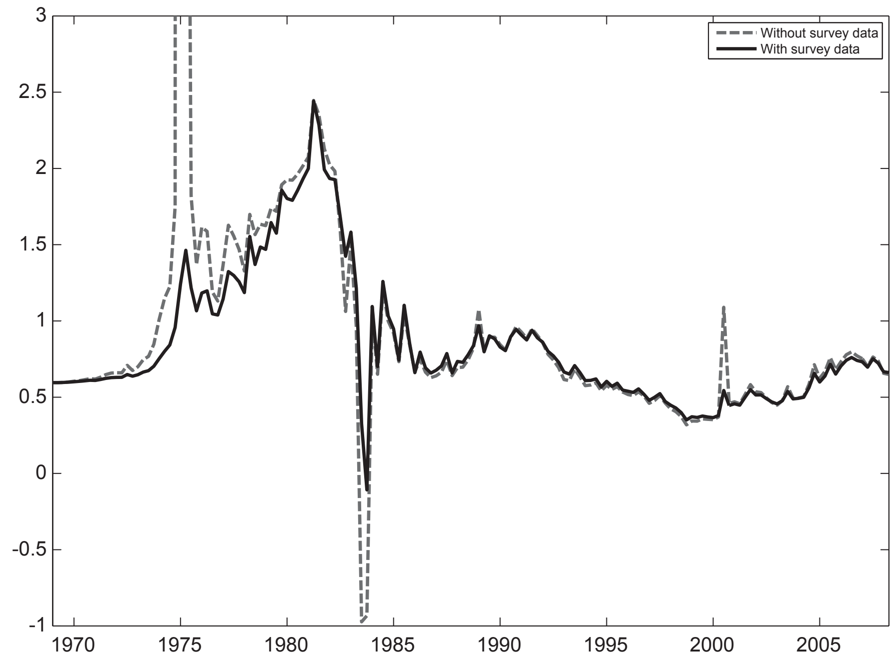


FIG. 6. Time Variation of Learning Agents' Perceived Inflation Target.

NOTES: Regression: $dlP_t = \beta_0^\pi + \beta_1^\pi dlP_{t-1}$. (Best-fitting model to SPF inflation survey.) Perceived inflation target $\beta_0^\pi/(1-\beta_1^\pi)$.

A great advantage of using survey data in the estimations is that it helps to avoid unrealistic jumps in model-implied inflation expectations. This is more pronounced for learning than for RE (see Figure 3). The estimated coefficients of CG-LS inflation expectations are smoother when survey data are used as an observable (Figure 4), which translates into smoother perceived long-run inflation targets as well (Figure 6). Interestingly, the announcement of the Volker disinflation did not decrease expectations about long-run inflation right away. Erceg and Levin (2003) explains this with imperfect information: agents were not able to tell whether increased interest rates mean the central bank had changed its inflation target or whether it was just a temporary shock. Our results show that CG-LS learning is consistent with this story as well: it takes time for agents to revise their perceived long-run inflation target, because they need enough new data to revise their estimates.

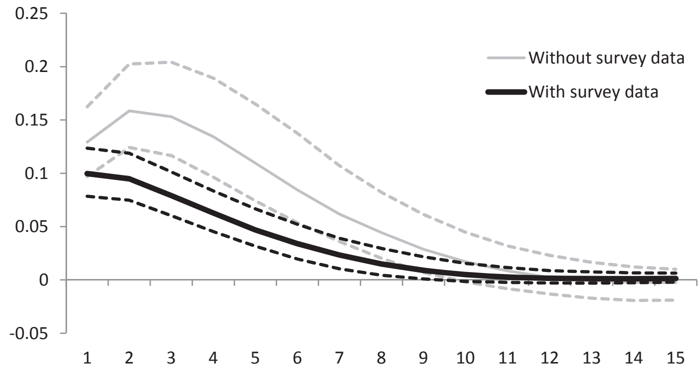
Posterior estimates and impulse responses. A general result in the literature is that adaptive learning enhances the propagation mechanism in DSGE models, but there is disagreement about which nominal frictions are important to match the dynamics of inflation. Milani (2007) finds that the structural inertia is systematically reduced when learning is introduced, while Slobodyan and Wouters (2012) find that this result is not general. We add to this debate by using the extra information in survey expectations of inflation. In this section, we report how the posterior estimates of those parameters that directly affect price stickiness change.[23] First, let us discuss the RE estimation results. Comparing columns (1) and (3) in Table 4, we can see that there is extra information in the inflation survey data as to which propagation channel is important. Column (1) presents estimates of the RE model without survey data, which is the benchmark estimate of SW. Column (3) shows that including inflation survey in the data set reduces the importance of price indexation (from a posterior median of 0.327 to 0.052) and wage stickiness (from 0.554 to 0.468). These parameter changes affect the inflation impulse responses (Figure 7). The reduction of wage stickiness causes a less persistent inflation response to the wage markup shock (the stochastic properties of the wage mark-up shock itself do not change; see the online appendix). Interestingly, despite the large reduction in the degree of inflation indexation, inflation's impulse response to the price markup shock does not change. This happens because the autocorrelation coefficient of the price markup shock increases (from a posterior median of 0.448 to 0.726) and this counteracts the effect of the smaller price indexation.[24]

Using survey data in the learning model does not change the parameters, the only significant parameter change being in the gain parameter of inflation expectations (compare columns (2) to (4)). The posterior mean of the constant gain parameter is high in both estimations, but the gain parameter decreases from 0.188 to 0.141 when survey data are included in the data set. As we discuss in Section 3.2, this removes some unrealistic jumps in the evolution of beliefs. Since the model-implied inflation expectations fit the survey data well (see Figure 3), using survey data in the estimations does not systematically alter other structural parameters. In other words, modeling inflation expectations with this benchmark model or using survey data as an observable contains similar information about the structural parameters.

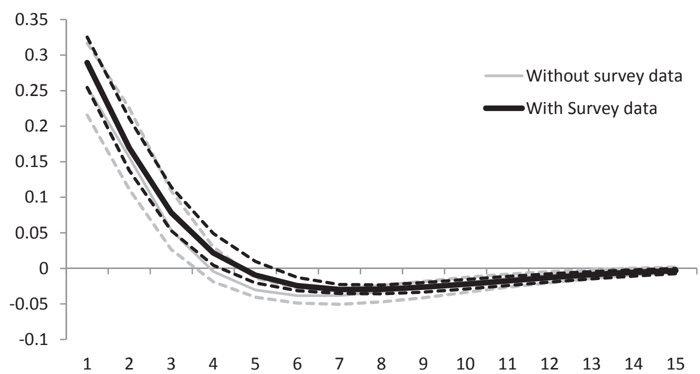
The impulse response of inflation under learning is a function of the time-varying CG-LS coefficients; therefore, it is time varying as well (Figure 8). Using survey data as an observable makes the CG-LS coefficients less volatile; therefore, the impulse responses of inflation are also more stable. Inflation impulse responses are stronger in the 1970s than in later years. For example, unexpected monetary policy shocks had a stronger destabilizing effect on inflation during the 1970s than after (Figure 8(c)). Similar results were found in Boivin and Giannoni (2006): inflation responds more strongly to unexpected changes in the interest rate before 1979 than after.

23. Other parameter estimates are so close both with and without survey data that we only report them in the online appendix. These parameter estimates do not change much probably because expectations (other than inflation) are modeled as consistent with the RE equilibrium. Estimation results of Slobodyan and Wouters (2009) show the same insensitivity of parameter estimates when learning expectations are consistent with the RE equilibrium.

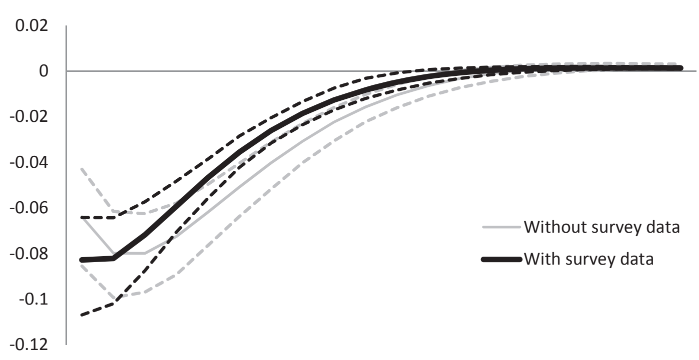
24. We find that impulse responses of other endogenous variables are not affected in estimations with survey data; therefore, we do not report them here. In the online appendix we report the variance–covariance analysis of inflation. We find that the relative importance of different shocks do not depend on whether the estimation is done with survey data, but depends on how the learning algorithm of inflation expectations is formulated (see Section 4).



(a) Wage mark-up shock



(b) Price mark-up shock

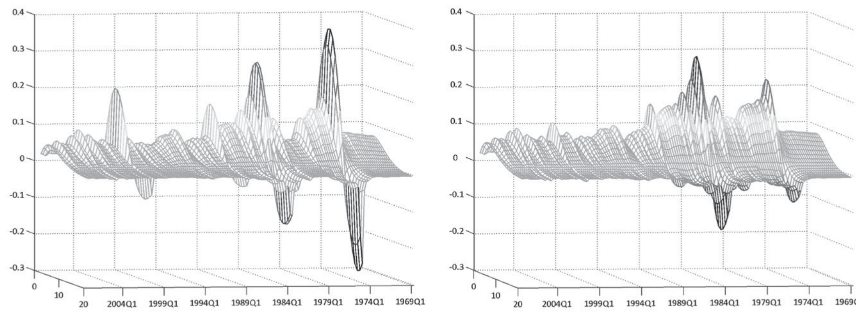


(c) Monetary policy shock

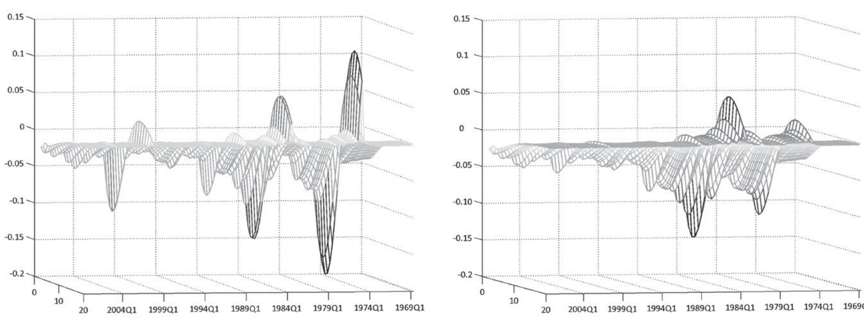
Fig. 7. Impulse Responses of Inflation, RE Model.

Notes: The plot shows impulse responses of inflation to price and wage markup shocks in the RE model (data set with and without inflation survey).

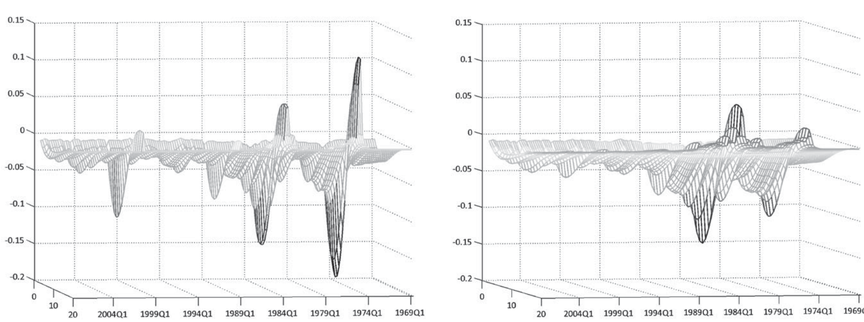
Data set without inflation survey   Data set with inflation survey



(a) Wage markup shock



(b) Productivity shock



(c) Monetary policy shock

FIG. 8. Impulse Responses of Inflation, Learning Model.

NOTES: The plot shows impulse responses of inflation to price and wage markup shocks in the RE model (data set with and without inflation survey). Impulse responses are a function of the time-varying learning coefficients.

An interesting difference between learning and RE impulse responses is that inflation under learning has a hump-shaped impulse response to a monetary shock. As learning is backward looking, it indirectly introduces lagged inflation into the Phillips curve, which has been shown to yield a hump-shaped impulse response function for

TABLE 6

POSTERIOR DISTRIBUTION STATISTICS

		(1)		(2)		(3)		(4)		(5)	
		dlP (benchmark)		*dlP dlCons*		*dlP lHours lHours*		*dlP dlCons*		*dlP dlGDP*	
Parameter	Symbol	Median	Std.	Median	Std.	Median	Std.	Median	Std.	Median	Std.
Wage stickiness	ξ_w	0.563	0.049	0.555	0.048	0.548	0.046	0.548	0.046	0.551	0.048
Price stickiness	ξ_p	0.480	0.035	0.446	0.033	0.467	0.037	0.453	0.035	0.462	0.037
Wage indexation	ι_w	0.319	0.107	0.314	0.101	0.342	0.110	0.326	0.106	0.319	0.109
Price indexation	ι_p	0.515	0.119	0.584	0.118	0.553	0.118	0.618	0.112	0.674	0.101
TR: inflation	r_π	1.398	0.104	1.432	0.111	1.404	0.121	1.468	0.111	1.423	0.115
TR: lag interest rate	ρ_R	0.777	0.029	0.775	0.029	0.767	0.029	0.778	0.023	0.776	0.031
TR: change in output	$r_{\Delta y}$	0.210	0.047	0.210	0.045	0.203	0.046	0.214	0.044	0.206	0.046
Aut. price markup shock	ρ_p	0.173	0.087	0.180	0.083	0.168	0.083	0.175	0.080	0.124	0.062
Std. price markup shock	σ_p	0.204	0.014	0.221	0.014	0.206	0.015	0.223	0.013	0.221	0.014
Gain—inflation	g^π	0.141	0.009	0.132	0.006	0.137	0.007	0.130	0.005	0.140	0.008
Gain—others	$g^{\mathrm{non}\pi}$	0.019	0.031	0.023	0.022	0.028	0.049	0.019	0.021	0.019	0.032
Measurement exp error	$\sigma_{\exp}$	0.176	0.010	0.177	0.011	0.175	0.009	0.173	0.009	0.184	0.011
Log. mg. likelihood		45.2		38.6		43.3		40.4		24.1	

NOTES: This table shows the median and standard deviation respectively, of the posterior distributions of the parameters most closely related to inflation dynamics. (TR= coefficients of the Taylor rule.)

inflation. The impulse response function of inflation under RE is not hump shaped, because of the smaller coefficient of price indexation.

In sum, our results are closer to Slobodyan and Wouters (2012) than to Milani (2007). We find that not all structural frictions decrease and price indexation even becomes more important under learning (compare columns (1) and (2) in Table 4). The importance of different frictions seems to be sensitive to the modeling choice of learning. We therefore believe that it is really important to reduce our degrees of freedom in the choice of learning algorithms. We believe our method is a good way of doing this: using survey data to choose the learning algorithm.[25]

4. ROBUSTNESS EXERCISES

In this section, we show that our results are robust to changing the small forecasting model of CG-LS inflation expectations and to changing the prior.

We redo our estimations using survey data, assuming that inflation expectations are formed with constant-gain learning with alternate forecasting models in Table 2. The results in Table 6 show that these alternate learning models still have much higher likelihood than the RE model, although the likelihood is somewhat smaller than in our benchmark model. Changing the small forecasting model of inflation expectations barely changes the posterior parameter estimates. To give an example, the median value of the posterior distribution of the price indexation is not significantly different from 0.6 either in the benchmark estimation (Table 4) or in the robustness estimations (Table 6). Likewise, the gain parameters do not change significantly compared to the benchmark estimation.

Our results are also robust to changing the priors. We reestimate the models with loose uniform prior distributions and find that the CG-LS model outperforms RE when survey data are used among the observables. Parameter estimates are also robust to changing the priors and price indexation remains low in the RE model when survey data are among the observables (see the online appendix).

5. CONCLUSION

The message of our paper is that using survey data of inflation provides useful information not present in macro data as to how to model private expectations. We contrast RE with adaptive learning and find that the data favor adaptive learning, but only when learning is modeled with a a small forecasting model. A small learning forecasting model conditions on a smaller information set than RE; therefore, the model dynamics are also different from those of the RE model. We show that this

25. The parameter estimates also depend on whether the DSGE model is small scale (like Milani 2007) or medium scale (e.g., Slobodyan and Wouters 2012 and our model).

simultaneously explains the evolution of the SPF inflation survey and the dynamics of U.S. macro data. A further advantage of using inflation survey data as one of the observables is that it helps to prevent unrealistic jumps in model-implied inflation expectations.

Combining inflation survey data with DSGE models opens up a new avenue for examining expectation formation. It is possible to examine other information contained in surveys or a wider set of assumptions about expectations. For example, one can use surveys on other macroeconomic variables, or use individual surveys to exploit more information than the mean, and one could combine surveys about the expectations of the central bank with those of the private sector in a monetary model. Another promising extension to our research would be to impose rationality bounds on adaptive learning as suggested by Marcet and Nicolini (2003). We leave these interesting extensions for future research.

LITERATURE CITED

Benveniste, Albert, Michel Métivier, and Pierre Priouret. (1990) *Adaptive Algorithms and Stochastic Approximations*. Berlin: Springer.

Binder, Michael, and Hashem Pesaran. (1997) "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation." *Econometric Theory*, 13, 877–88.

Blanchard, Olivier Jean, and Charles M. Kahn. (1980) "The Solution of Linear Difference Models under Rational Expectations." *Econometrica*, 48, 1305–11.

Boivin, Jean, and Marc P. Giannoni. (2006) "Has Monetary Policy Become More Effective?" *Review of Economics and Statistics*, 88, 445–62.

Branch, William A., and George W. Evans. (2006) "A Simple Recursive Forecasting Model." *Economics Letters*, 91, 158–66.

Bray, Margaret M. (1982) "Learning, Estimation, and the Stability of Rational Expectations." *Journal of Economic Theory*, 26, 318–39.

Calvo, Guillermo A. (1983) "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics*, 12, 383–98.

Canova, Fabio, and Luca Gambetti. (2010) "Do Expectations Matter? The Great Moderation Revisited." *American Economic Journal: Macroeconomics*, 2, 183–205.

Carboni, Giacomo, and Martin Ellison. (2009) "The Great Inflation and the Greenbook." *Journal of Monetary Economics*, 56, 831–41.

Christiano, Lawrence J. (2002) "Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients." *Computational Economics*, 20, 21–55.

Cogley, Timothy, and Thomas J. Sargent. (2002) "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual 2001*, Vol. 16, pp. 331–88. Cambridge, MA: National Bureau of Economic Research.

Coibion, Olivier, and Yuriy Gorodnichenko. (2010) "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." NBER Working Paper No. 16537.

Del Negro, Marco, and Frank Schorfheide. (2008) "Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities)." *Journal of Monetary Economics*, 55, 1191–1208.

Del, Negro, Marco, and Stefano Eusepi. (2011) "Fitting Observed Inflation Expectations." *Journal of Economic Dynamics and Control*, 35, 2105–31.

Erceg, Christopher J., and Andrew T. Levin. (2003) "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics* 50, 915–44.

Evans, George W., and Seppo Honkapohja. (2001) *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Ireland, Peter N. (2003) "Irrational Expectations and Econometric Practice: Discussion of Orphanides and Williams,' Inflation Scares and Forecast-Based Monetary Policy.'" Working Paper 2003–22, Federal Reserve Bank of Atlanta.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. (2004) "Disagreement about Inflation Expectations." In *NBER Macroeconomics Annual 2003*, Vol. 18, pp. 209–70. Cambridge, MA: National Bureau of Economic Research.

Marcet, Albert, and Juan Pablo Nicolini. (2003) "Recurrent Hyperinflations and Learning." *American Economic Review*, 93, 1476–98.

Marcet, Albert, and Thomas J. Sargent. (1989) "Convergence of Least Squares Learning Mechanisms in Self-referential Linear Stochastic Models." *Journal of Economic Theory*, 48, 337–68.

Milani, Fabio. (2007) "Expectations, Learning and Macroeconomic Persistence." *Journal of Monetary Economics*, 54, 2065–82.

Milani, Fabio. (2011) "Expectation Shocks and Learning as Drivers of the Business Cycle." *The Economic Journal*, 121, 379–401.

Orphanides, Athanasios, and John C. Williams. (2005) "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control*, 29, 1927–50.

Sargent, Thomas J. (1999) *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press.

Sargent, Thomas J., Noah Williams, and Tao Zha. (2004) "Shocks and Government Beliefs: The Rise and Fall of American Inflation." NBER Working Paper No. 10764.

Sims, Christopher A. (2002) "Solving Linear Rational Expectations Models." *Computational Economics*, 20, 1–20.

Slobodyan, Sergey, and Raf Wouters. (2009) "Estimating a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models." 2009 Meeting Papers 654, Society for Economic Dynamics.

Slobodyan, Sergey. (2012) "Learning in an Estimated Medium-Scale DSGE Model." *Journal of Economic Dynamics and Control*, 36, 26–46.

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

Uhlig, Harald. (1999) "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily." In *Computational Methods for the Study of Dynamic Economies*, edtied by R. Marimon and A. Scott, pp. 30-61. Oxford, UK: Oxford University Press.