Simple Rules in Estimated DSGE Models with Financial Frictions

Yuliya Rychalovskaya and Sergey Slobodyan*

University of Namur

NRU HSE St Petersburg and CERGE-EI

preliminary draft (please do not circulate without permission)

April 25, 2017

Abstract

We estimate a Smets-Wouters type model with financial frictions where the adaptively learning agents act as econometricians and update their beliefs using past information. The agents are allowed to select among several simple rules to predict every forward-looking variable of the model. They use Bayesian Model Averaging techniques to determine the weights which are used to aggregate individual forecasting rules' predictions into a common forecast. We estimate the resulting models and discuss the effects different sets of allowed forecasting rules have on the results. While introducing certain simpler rules into the allowed rules pool could lead to a better overall model fit, the baseline, more complex forecasting rule, always dominates the agents' forecasts.

JEL Classification: C5, D84, E3

Keywords: Inflation, Expectations, Adaptive Learning, DSGE models, model averaging

*This work is supported by the EU 7th framework collaborative project "Integrated Macro-Financial Modelling for Robust Policy Design (MACFINROBODS)" grant no. 612796.

1 Introduction

The expectation formation mechanism remains one of the least explored issues in the vast literature on estimated DSGE models. Mostly, empirical DSGE models employ the Rational Expectations (RE) assumption, which endows the agents with complete information about the model and all parameters, including data-generating processes of the stochastic processes. Their forecasts of the forward-looking variables are therefore model consistent. The assumption of model consistency could be thought unrealistic, especially given rapidly changing structure of the economy, consequences of the recent financial crisis which forced the Central Banks around the world to adopt non-traditional monetary policy measures with highly uncertain impact on the economy, and a possible period of low future rate of productivity growth.

There has been a long tradition in the economic literature of modeling economic agents that forecast the future based on incomplete information, and have to perform actions informed by such forecasts. One option that such agents could use is learning the underlying economic structure in real time as more data becomes available. As such agents behave as real life econometricians, one could presume that this type of modeling the agents' forecasting process is more realistic and might enable more reasonable conclusions to be drawn about the way the public's expectations are formed, and the effect of these expectations on the economic outcomes.

In this paper we follow Rychalovska, Slobodyan, and Wouters (2015) who incorporated adaptively formed expectations (Evans and Honkapohja, 2001, Milani, 2007 and Orphanides and Williams, 2007) into a version of the Smets and Wouters (2007) DSGE model with a financial accelerator (FA) as in Bernanke et al. (1999). As is standard in the adaptive learning literature, the agents are assumed to know the structure but not the parameters of the model. They learn them by formulating linear econometric models with coefficients based on their economic perceptions (beliefs), and re-estimating these models as soon as new information arrives to obtain new beliefs.

Following Slobodyan and Wouters (2012b), we assume that forecasting models are re-estimated every period on the basis of the Kalman filter learning algorithm. While Rychalovska et al. (2015) assume that agents' forecasts are based on a very small forecasting model, where expected value of a forward-looking variable depends on a constant and two lags of this variable, here we investigate a broader set of assumptions regarding the agents' expectations formation mechanism. In particular, we allow them to entertain a set of different forecasting models, and to use as a forecast a linear combination of these models' predictions, using either fixed weights or the weights that depend on past forecasting performance of the models. These forecasting models remain small. Their particular choice is informed by a series of papers by Cars Hommes and co-authors (cf., Anufriev and Hommes 2012, Hommes 2011, and Bao, Hommes, and Makarewitz 2016), who show that the behavior of subjects in learn-to-forecast experiments is well described by the assumption that people follow a very small set of simple forecasting rules with fixed coefficients and occasionally switch the rules on the basis of their past forecasting performance. As our previous work has indicated that the most important forecasting rule in Smets and Wouters (2007) type models is the inflation one, we concentrate on varying only the agents' inflation forecasting and keep the forecasting equations for the remaining forward-looking variables fixed.

The model is estimated using Bayesian methods. On the basis of the estimation and simulation results, we assess how adaptive learning (AL) algorithm can modify the transmission mechanism of the model, investigate ability of AL to generate additional macroeconomic fluctuations in line with real data, and study whether mixing different forecasting rules for inflation is beneficial for the overall model fit and its forecasting performance.

The rest of the paper is organized as follows: in Section 2, we briefly present the model of Rychalovska et al. (2015); Section 3 contains the estimation methodology, discussion of forecasting rules used, and results; Section 4 describes the evolution of the agents' beliefs when different combinations of forecasting rules are used, and investigates whether the set of rules and the characteristics of these rules (initial beliefs, fixed or time varying weights) are materially affecting the contribution of the financial frictions to the transmission mechanism in the model with AL. The Section 5 concludes.

2 The model with a financial accelerator under learning

The FA framework, which implies that credit markets amplify and propagate shocks to the real economy, is widely implemented and analyzed in the post-crisis macroeconomic literature. The approach used in Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), and Cespedes et al. (2004), among others, captures firms' balance sheet effects on investment by relying on a one-period stochastic optimal debt contract with costly state verification, which leads to endogenous interest rate spread. Bernanke et al. (1999) introduce the agency problem with asymmetric information in order to model this difference between the cost of external sources of funding and the opportunity cost of funds internal to the firm. Due to the agency problem in lending, the external finance premium depends inversely on the borrowers' net wealth and thus is countercyclical, amplifying the effects of monetary and financial shocks. Below, we present log-linearized equations of the model of Rychalovska et al. (2015), which estimates a Smets and Wouters (2007) type model with Bernanke et al. (1999) financial accelerator under assumption that the agents are adaptive learners.[1]

The model economy consists of households, final and intermediate goods producers, a monetary authority and a financial sector. Intermediate-sector firms are monopolistically competitive. They produce differentiated goods, decide on labor and capital input and set prices according to the Calvo (1983)

[1] Several papers performed a similar exercise under RE, cf., Chtistiano et al. (2010), De Graeve (2008) and Christensen and Dib (2008).

model. Households supply homogenous labor to an intermediate labor union, which differentiates the labor services and sets wage rates. The prices that are not optimised are partially indexed to past inflation rates. The log-linearized model equations of the households and production sector are summarized in the Appendix.

The financial sector contains capital good producers, financial intermedies and entrepreneurs. Capital producers accumulate new capital and sell it to entrepreneurs. Entrepreneurs rent capital stock to intermediate firms and borrow from the bank in order to finance capital acquisitions. The competitive bank obtains resources for lending by issuing deposits to households. As in Bernanke et al. (1999), financial contracts issued to households pay a real return which is not contingent upon the realization of the shock. The financial intermediation between the banks and entrepreneurs is subject to friction based on the agency problem. In particular, after the purchase of the capital stock, each entrepreneur receives an idiosyncratic productivity shock that affects their capital holdings and the return on capital holdings. Banks have to pay a monitoring cost to infer the realized return. Due to the presence of such an asymmetric information problem, the optimal financial contract, which maximizes the entrepreneur's payoff subject to the required rate of return of lenders, implies the existence of an endogenous external finance premium over the riskless rate, which entrepreneur has to pay in order to borrow funds. The external finance premium is inversely related to the entrepreneur's leverage ratio.[2]

2.1 Financial sector

In this section, we present the log-linearized equations describing the behavior of capital-good producers and entrepreneurs. For more detailed description of the financial block micro-foundations, we would like to refer to the original Bernanke et al. (1999) model as well as to the paper of Rychalovska et al. (2015).

Competitive *capital-goods producers*, owned by households, combine investment goods, purchased from the final good producers, with the existing capital stock, rented from the entrepreneurs, to produce new capital goods, K_{t+1}, which are sold to entrepreneurs at price Q_t. The optimization problem of capital-goods producers consists of choosing the level of investment I_t to maximize the real expected profits, and after detrending and log-linearization becomes

$$\widehat{i}_t = \frac{1}{(1+\overline{\beta}\gamma)}(\widehat{i}_{t-1} + (\overline{\beta}\gamma)\widehat{i}_{t+1} + \frac{1}{\gamma^2 S''}\widehat{Q}_t \ ) + \widehat{q}_t, \tag{1}$$

where S'' is the steady-state elasticity of the capital adjustment cost function, $\overline{\beta}$ is technology growth rate (γ) adjusted discount factor, and $\widehat{q}_t$ a disturbance to

[2] The equilibrium condition on financial markets is derived from the optimal-debt contract problem, which maximizes the welfare of the entrepreneur, combined with the zero-profit condition of the bank. The details of the financial contract specification and derivations can be found in Appendix A of Bernanke et al. (1999).

the investment–specific technology process assumed to follow a AR(1) process with an iid–normal error term: $\widehat{q}_t = \rho_q \widehat{q}_{t-1} + \epsilon^i_t$. Higher elasticity of the cost of adjusting capital reduces the sensitivity of investment $(\widehat{i}_t)$ to the real value of the existing capital stock $(\widehat{Q}_t)$. The evolution of the capital stock is then represented by the following expression:

$$\widehat{k}_t = (1 - \frac{i_*}{\overline{k}_*})\, \widehat{k}_{t-1} + \frac{i_*}{\overline{k}_*}\widehat{i}_t + \frac{i_*}{\overline{k}_*}(1 + \overline{\beta}\gamma)\gamma^2 S'' \widehat{q}_t, \tag{2}$$

where symbols with asterics represent the steady-state values.

Entrepreneurs, who are risk neutral and survive until the next period with probability $\varkappa$, use their own funds (the net worth, N_{t+1}) and loans from the bank (B_{t+1}) to finance capital that is rented to the production sector. At the end of period t, entrepreneurs purchase capital K_{t+1} from capital-goods producers at price Q_t and borrow $B_{t+1} = Q_t K_{t+1} - N_{t+1}$. After observing the idiosyncratic productivity shock ω_{t+1}, the entrepreneur decide on the capital utilization (U_{t+1}). Log-linearized optimality condition to the entrepreneur's decision problem is

$$\widehat{u}_t = ((1-\psi)/\psi)\widehat{r}^k_t, \tag{3}$$

with ψ being the elasticity of the capital utilization cost function and r^k_{t+1} is the rental rate for capital services to intermediate-goods firms. The log linearized relation for capital services is given by

$$\hat{k}^S_{t+1} = \widehat{u}_{t+1} + \widehat{k}_{t+1}. \tag{4}$$

The average (aggregated over all the entrepreneurs) rate of return on capital purchased at time t is given by

$$E_t \widehat{R}^K_{t+1} = \frac{1-\tau}{\overline{R}^K} E_t \widehat{Q}_{t+1} + \frac{\overline{r}^k}{\overline{R}^K} E_t \widehat{r}^k_{t+1} - \widehat{Q}_t, \tag{5}$$

where $\overline{R}^K$ denotes the steady-state return to capital, $\overline{r}^k$ is the steady-state rental rate, and τ the depreciation rate.

The log-linearized optimality condition, which determines the link between the external financing costs, capital purchases and entrepreneurial financial position, is given by

$$\widehat{prem}_t = E_t \hat{R}^K_{t+1} - (\hat{R}^n_t - E_t[\widehat{\pi}_{t+1}] + \hat{\varepsilon}^b_t) = -el \left\{ E_t \left[\widehat{N}_{t+1} - \widehat{Q}_t - \widehat{k}_{t+1} \right] \right\}, \tag{6}$$

which indicates that the cost of external financing is composed of the premium for borrowed external funds, the risk-free interest rate and an exogenous disturbance $\hat{\varepsilon}^b_t$ that describes fluctuations in the risk premium not captured by the financial frictions of Bernanke et al. (1999). Parameter el represents the elasticity of the external finance premium to the expected change in the financial conditions.

The log-linearization of the entrepreneurial net worth is given by the following accumulation equation:

$$\widehat{N}_{t+1} = \varkappa \overline{R}^K \left[\frac{\overline{K}}{\overline{N}} \left(\hat{R}_t^K - E_{t-1}\hat{R}_t^K \right) + E_{t-1}\hat{R}_t^K + \widehat{N}_t \right] + \varepsilon_t^{nw}, \tag{7}$$

where $\overline{K}/\overline{N}$ is the steady-state ratio of capital to net worth, i.e. the inverse of the leverage ratio, and ε_t^{nw} is the shock to the net worth, which follows a stationary AR(1) process. Equation (7) demonstrates that, in general terms, the endogenous variations in net worth in the next period come from unexpected changes in the real return on capital, which might come from the shocks that reduce the rental rate of capital or the market value of capital. As a result of such shocks, the entrepreneurial net worth will drop, leading to a fall in investment. Because both the rental rate and the price of capital are forward-looking variables, learning could affect investment through the FA channel. The net worth can be also expressed as a function of the risk-free interest rate and the exogenous and endogenous finance premia:

$$\widehat{N}_{t+1} = \varkappa \overline{R}^K \left[\frac{\overline{K}}{\overline{N}} \widehat{R}_t^K - \left(\frac{\overline{K}}{\overline{N}} - 1 \right) \left(\widehat{R}_{t-1} + \widehat{\varepsilon_{t-1}^b} \right) - el \left(\frac{\overline{K}}{\overline{N}} - 1 \right) \left(\widehat{k}_t + \widehat{Q}_{t-1} - \widehat{N}_t \right) + \widehat{N}_t \right] + \widehat{\varepsilon}_t^{nw}. \tag{8}$$

The values of the parameters $\varkappa$, $\overline{K}/\overline{N}$ and el determine the impact of financial frictions on the real economy. The higher the entrepreneurial survival rate and the capital to the net worth steady-state ratio, the more persistent the evolution of net worth will be. Combined with the higher elasticity of the external finance premium, this would imply a stronger response of the wedge between the expected return on capital and the risk-free rate. Therefore, shocks affecting entrepreneurial net worth would have greater real effects.

3 Introducing adaptive learning

We depart from the assumption of RE and assume that agents possess incomplete knowledge about the economic environment (model structure and parameters). They form beliefs about future evolution of the forward-looking variables on the basis of the information they observe. As in Marcet and Sargent (1989) and Evans and Honapohja (2001), agents gradually learn the "true" parameters of the model by using a learning algorithm. In this section, we present a general description of a Kalman filter learning setup[3]. For more details on the description and implementation of this learning process, see Slobodyan and Wouters (2012b).

[3] The alternative, widely used, learning method is the constant gain Recursive Least Squares (RLS). Sargent and Williams (2005) demonstrate that both learning methods mentioned above are asymptotically equivalent on average. However, their transitory behavior may differ significantly. In particular, the Kalman filter tends to result in a faster adjustment of agents' beliefs due to the fact that the learning gain is not assumed to be constant.

In general form, the RE solution of the DSGE model is given by the following expression:

$$\begin{bmatrix} y_t \\ w_t \end{bmatrix} = \mu + T \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + R\epsilon_t, \tag{9}$$

where y_t is the vector of endogenous variables of the model w_t is an exogenous AR(1) process. T and R are time-invariant non-linear functions of the model structural parameters Θ. Under RE for log-linearized models, the intercept μ is normally a vector of zeros [4]. There are 8 exogenous structural shocks and 2 measurement errors. The model variables summarized by vector y can be grouped into the state variables y^s that appear with a lag, forward variables y^f that appear with a lead, and the so-called static variables[5]. Agents have to forecast 7 forward-looking variables: consumption, investment, hours worked, price and wage inflations, return on capital and asset prices. Deviation from the RE assumption implies that agents do not have sufficient information about the model and its parameters to formulate the REE-based expectations, $E_t y_{t+1}$. Thus, they cannot derive the law of motion (9). To obtain the solution, they have to formulate the so-called Perceived Law of Motion (PLM), which relates the value of the forward variable j to the model state variables X_j using a reduced-form linear function:

$$y^f_{j,t} = \beta_{j,t-1} X_{j,t-1} + u_{j,t}. \tag{10}$$

The forecasting model is then represented in the SURE format and composed of the variable-specific equations of the form (10). In this paper, in period t, agents predict only next-period variables which are present in Euler equations of firms and consumers, therefore implementing the *Euler equation learning* as in Evans and Honkapohja (2001) [6]. The error term $u_{j,t}$ in (10) consists of linear combination of the true model errors ϵ_t and has a variance-covariance matrix Σ. Matrix X_j includes a set of variables that are used to form predictions about forward-looking variable j. Thus, X_j may consist of all the state variables of the model y^s or a subset of y^s. In addition, the vector of beliefs β_j may also include the constant which would mean that agents make inferences not only about the persistence of economic processes but also about growth rates. Finally, the information set X_j can be significantly more restrictive than the vector of all the state variables y^s, which would be used under RE.

The Kalman filter learning algorithm consists of the following steps[7]:

1. Formulation of the initial beliefs. Following Slobodyan and Wouters (2012a and 2012b), we apply the standard assumption in the learning literature and assume that initial beliefs are consistent with the REE.

2. Update. The Kalman filter is used to obtain the updated estimates of the vector of beliefs β and the error covariance matrix

[4] It can be a nonzero vector only for observable variables that are not detrended.

[5] The sets y^f and y^s could intersect.

[6] An alternative type of learning - long-horizon learning, advocated by B. Preston (2005)- implies that agents forecst economic variables infinitely many periods ahead.

[7] More technical details can be found in Slobodyan and Wouters (2012b).

3. Prediction.

The best estimates of the beliefs $\beta_{t|t-1}$ generated in the prediction step of the Kalman filter are then used to calculate subjective expectations $\hat{E}_t y_{t+1}$ of forward–looking variables according to forecasting equations (10). Substituting these expectations into the structural representation of the model, we obtain the Actual Law of Motion (ALM) of the system in a purely backward–looking form:

$$\begin{bmatrix} y_t \\ w_t \end{bmatrix} = \mu_t + T_t \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + R_t \epsilon_t. \tag{11}$$

Introducing AL does not affect the initial steady state of the system, and at time $t = 0$ we start from the RE equilibrium solution given by equation (9).

Typically, we let our agents use only a very limited set of variables in their forecasting equations. This set does not include exogenous processes. Learning agents must make inferences about the shocks hitting the economy having data only on observed variables, which leads to a typical incomplete information problem. Differently from incomplete information problems, our agents update their beliefs over time, which results in time variation in the way innovations in observables are allocated to different shocks. Thus, not only is the transmission mechanism of the model under AL with limited forecasting rules generically different from that under RE, it is also time varying. Both of these features of adaptive learning play significant role in the ability of a model to explain the data, which we will discuss in more details below.

As mentioned previously, the baseline model of Rychalovska et al. (2015) is estimated assuming that the agents' PLM for every forward looking variable contains only two lags of itself and a constant, plus some additional variables specific to a particular forward-looking variable. In particular, inflation PLM is AR(2) with a constant plus marginal cost, where all coefficients are time-varying due to the beliefs updating. Bao et al. (2012) run an experiment in which subjects were asked to forecast next period price in a simple learning-to-forecast environment. They have shown that the whole variety of experimental outcomes could be explained by subjects using a few very simple forecasting rules, occasionally changing the one actually used on the basis of the past forecast performance of the rule. We use the set of forecasting models described by Bao et al. (2012) in a model selection exercise, taking into account that there is no particular reason for preferring exactly the baseline forecasting rule. As it was demonstrated previously (cf., Slobodyan and Wouters 2012b, 2017) that it is the inflation PLM which is the most important for the model's overall fit, we allow the agents to vary only the inflation PLM. The additional PLMs then include the following ones: just the constant, 'lagged' AR(2) with constant, and 'lagged' AR(2) with running average inflation.[8] We allow agents to combine these forecasts either by simple equal weights (EW version) or by a Bayesian

[8] To be precise, these forecasting rules generate $\hat{E}_t \pi_{t+1}$ as follows:

Baseline	π_t, π_{t-1}, mc_t
Constant	π_t^{av}
Lagged AR(2)	π_{t-1}, π_{t-2}
Lagged AR(2) with running average	π_{t-1}, π_{t-2}, π_t^{av}

Model Averaging method in which the weight attached to each model varies with the past forecasting performance of each of the small forecasting models (BIC version)[9].

4 Results

4.1 Data and measurement equations

The model is estimated using the standard set of seven Smets and Wouters (2003, 2007) observables — per capita real GDP, consumption, investment, and wage growth rates, hours worked, GDP deflator inflation, federal funds rate, — plus stock prices index and the credit spread[10] All the data are quarterly U.S. time series, with nominal variables deflated by the GDP deflator. We estimate the model for the sample period 1954:1 - 2014:2. The long data sample, including relatively calm periods both before and after the Great Inflation of the 70es and early 80es, is chosen in order to ascertain the effects of time varying beliefs vs. time varying weight of different forecasting rules, when the agents perform adaptive learning. The measurement equations of the model are presented below:

$$\begin{bmatrix} dlGdp_t \\ dlCons_t \\ dlInv_t \\ dlWage_t \\ lHours_t \\ dlP_t \\ FedFundsR_t \\ dSP500_t \\ Spread_t \end{bmatrix} = \begin{bmatrix} \overline{\gamma_y} \\ \overline{\gamma_c} \\ \overline{\gamma_i} \\ \overline{\gamma_w} \\ \overline{l} \\ \overline{\pi} \\ \overline{r} \end{bmatrix} + \begin{bmatrix} \widehat{y}_t - \widehat{y}_{t-1} \\ \widehat{c}_t - \widehat{c}_{t-1} \\ \widehat{i}_t - \widehat{i}_{t-1} \\ \widehat{w}_t - \widehat{w}_{t-1} \\ \widehat{l}_t \\ \widehat{\pi}_t \\ \widehat{R}_t^n \\ \widehat{N}_t - \widehat{N}_{t-1} + me_{N,t} \\ \widehat{prem}_t + me_{S,t} \end{bmatrix}, \tag{12}$$

where l and dl stand for log and log difference, respectively. Unlike Smets and Wouters (2007), the trends for output, consumption, investment and wages growth rates are estimate separately. $\overline{\pi} = 100(\Pi_* - 1)$ is the quarterly steady-state inflation rate and $\overline{r} = 100(\overline{\gamma}^{\sigma_c}\Pi_*/\beta - 1)$ is the steady–state nominal interest rate. Given the estimates of the average trend growth rate and the steady–state inflation rate, the latter will be determined by the estimated discount rate.

Here π_t^{av} is the weighted average of past inflation rates with the weights geometrically declining into the past, as in constant gain evaluation. We use parameter ρ to generate the weights.

[9] It is interesting to note that in our estimations, a simple fixed-weight forecast combination works better than sophisticated time-varying re-weighting of the forecasts. Timmerman (2006) surveys a broad literature which reaches a similar conclusion, namely that simple forecast combinations "often dominate more refined combination schemes aimed at estimating the theoretically optimal combination weights". For demonstration of dominance of simple averaging of forecasts of quarterly GDP growth, see Watson and Stock (2004).

[10] The stock price index is S&P500. The credit spread is defined as the difference between the corporate BAA yield and the AAA yield.

Finally, $\bar{l}$ refers to steady-state hours worked. Measurement errors denoted by $me_{N,t}$ and $me_{S,t}$ are added to the observation equations of financial variables.

4.2 Bayesian estimation under adaptive learning

The log-linearized versions of the models are estimated using Bayesian methods. These methods combine a likelihood function of the data with a prior density to derive the posterior distribution of the structural parameters. The prior density contains information about the model parameters from other sources (microeconometric and calibration evidence). The estimation procedure includes: first, the estimation of the mode of the posterior distribution by maximizing the log posterior function, and second, the Metropolis–Hastings algorithm, which is used to compute the posterior distribution and to evaluate the marginal likelihood of the model[11],[12]. For estimation purposes, the model is represented in the state-space form, which combines equation (11), the state equation, with the measurement equations (12). Given the state space representation of the model, the Kalman filer is used to evaluate the likelihood. For evaluating the model under imperfectly rational beliefs, we use the toolbox developed by Slobodyan and Wouters (2012b), which implements procedures for AL within the Dynare 4.25 Matlab toolbox. As described in Section 2, the estimation procedure under learning differs from the inferences under the RE due to time variation in beliefs (**??**), which implies the dynamic solution of μ_t, T_t and R_t[13]. The values of μ_t, T_t and R_t are used to form expectations of the next period model variables in the main Kalman filter step and are used to calculate the model likelihood. In other words, the standard Kalman filter incorporates an additional filtering stage used to update agents' beliefs.

In this paper, we estimate the AL model specification with relatively small information set (given by matrix X_j in (10)) used by agents to form their beliefs about the forward-looking variables. Generally, we assume that the forecasting equation (10) for every forward-looking variable includes two lags and a constant, but some contain additional variables[14], which ensure that the agents use the financial data in predicting real activity and vise versa.

The priors for the estimation are chosen following Smets and Wouters (2003, 2007). These papers present a detailed description of the estimation methodology as well as the justification for the choice of priors. The priors for additional parameters related to the financial frictions are based on calibration exercises and previous literature (Bernanke et al., 1999, Merola, 2013 and De Graeve, 2008). In particular, $\overline{R}^K$ is calibrated at 1.0129. Parameters $\varkappa$, $\overline{k}/\overline{N}$, and el are assumed to have Normal priors with sufficient standard deviations.

[11] For more details on Bayesian estimation of DGSE models, see An and Schorfheide (2007).

[12] At the moment, MCMC step hasn't been performed yet. All estimations presented are based on the posterior mode.

[13] These matrices do not change over time during the estimation under RE.

[14] PLM for investment includes net wealth, for inflation marginal cost, and for price of capital investment and return on capital.

4.3 Estimation results. Model fit

Table 1 reports the logarithms of posterior probability at the mode for the various estimated specifications. We compare the results for models with RE and AL. In addition to the baseline AL model with the set of forecasting rules described in the previous subsection, we present the results for various combinations of forecasting rules for inflation, keeping all other PLMs the same. The inflation PLMs, then, are as follows:

$$\begin{array}{ll} R_1\text{: Baseline} & \hat{E}_t\pi_{t+1} = \mu_1 + \beta_1^1\pi_t + \beta_2^1\pi_{t-1} + \beta_3^1 mc_t, \\ R_2\text{: AR(2) lagged} & \hat{E}_t\pi_{t+1} = \mu_2 + \beta_1^2\pi_{t-1} + \beta_2^2\pi_{t-2}, \\ R_3\text{: mean} & \hat{E}_t\pi_{t+1} = \mu_3, \\ R_4\text{: AR(2) lagged+}\pi^{av} & \hat{E}_t\pi_{t+1} = \mu_4 + \beta_1^4\pi_{t-1} + \beta_2^4\pi_{t-2} + \beta_3^4\pi_{t-1}^{av}. \end{array}$$

In an estimated model, the agents could then have up to four forecasting models. Each of these models also contains six forecasting rules for variables other than inflation. These rules never change. The forecasting rule for inflation, however, could be selected from the set $\{R_1, R_2, R_3, R_4\}$. With a slight abuse of notation, we also denote the whole forecasting model (six fixed forecasting rules plus $R_1 ... R_4$ rule for inflation) of the agents as $R_1 ... R_4$. The agents are then endowed with a different combination of the forecasting models, where we always keep R_1, and add the rest in all possible combinations.

Table 1 also contains results of estimations where forecasting models are combined with either fixed or time varying weights. When the weights are fixed, each forecasting model has a weight equal to $1/R$, where R is the number of forecasting models used in this particular estimation. For estimations with variable weights, the initial weights are taken to be $1/R$, while the weights evolution over the estimation sample is given by

$$\begin{aligned} w_i &= \frac{\exp\left(-0.5 \cdot (BIC_i - \min BIC_i)\right)}{\sum_i \exp\left(-0.5 \cdot (BIC_i - \min BIC_i)\right)}, \\ BIC_i &= t \cdot \ln\left(\left|\Sigma_i\right| / t\right) + df \cdot \ln t. \end{aligned}$$

Thus, the weights are determined by the BIC statistics of a given rule, which, in turn, depends on the second moments matrix of the forecasting errors produced by the model R_i. By including the term proportional to the number of degrees of freedom, df, BIC statistics penalizes large forecasting rules that contain many predictors.[15]

The estimation results suggest that the model with Kalman filter learning fits the data significantly better than the RE model: the log posterior difference of 40 signifies an overwhelming evidence in favor of the AL model. Therefore, RE hypothesis appears to be restrictive. This result is in line with the previous studies which analysed the performance of the Smets and Wouters type models

[15] In work of Hommes and co-authors on switching heuristics, change in the share of agents using different rules is gradual, while it seems to be immediate in our specification. However, our measure of past rule's performance, BIC, incorporates the whole past history, while fitness measures in switching heuristics tend to involve only very recent performance.

Table 1: Model Comparison in Terms of Marginal Likelihood

Model specification	*fixed weights*	*variable weights*
RE	-2019.30	
AL: R_1	-1979.71	
AL: R_1, R_2	-1969.09	-1980.59
AL: R_1, R_2, R_3	-1977.19	-1978.81
AL: R_1, R_2, R_3, R_4	-1970.05	-1978.40
AL: R_1, R_3	-1978.02	-1976.54
AL: R_1, R_3, R_4	-1977.79	-1978.78
AL: R_1, R_4	-1972.59	-1977.68
AL: R_1, R_2, R_4	-1969.69	-1981.22

without financial frictions under learning (Slobodyan and Wouters, 2012a and 2012b). Further, it is evident that there are combinations of the forecasting rules which are significantly better than just the baseline R_1: models with R_1 and R_2, R_1, R_2, and R_4, and all four forecasting models, all produce significantly better log posterior (further 10 units). As all these winning combinations involve the model R_2, lagged AR(2) specification in inflation, we investigate this particular forecasting rule further.

The forecasting rules R_1 and R_2 appear rather similar: while the former implies $\hat{E}_t\pi_{t+1} = \mu_1 + \beta_1^1\pi_t + \beta_2^2\pi_{t-1} + \beta_3^1 mc_t$, the latter gives $\hat{E}_t\pi_{t+1} = \mu_2 + \beta_1^2\pi_{t-1} + \beta_2^2\pi_{t-2}$. However, there are two significant differences between the two rules: no marginal cost in R_2, and the AR(2) component is lagged by one period in R_2 relative to R_1. This suggests that perhaps the forecasting rule R_1 is mis-specified, and simply adding one more lag of inflation could give us a model with very good model fit. However, this turns out to not be the case, as the model R_5 in which the inflation PLM is given by $\hat{E}_t\pi_{t+1} = \mu_5 + \beta_1^5\pi_t + \beta_2^5\pi_{t-1} + \beta_3^5\pi_{t-2} + \beta_4^5 mc_t$ is producing log likelihood that is even worse than in the baseline, -1981.38. Combining forecasting models R_1 and R_5 only returns us approximately back to the baseline model's fit, -1979.29.

Let us look at the time evolution of the agents' beliefs in forecasting rules R_1 and R_2. Figures 1-3 present the evolution of beliefs $\left\{\mu_1, \beta_1^1, \beta_2^1, \beta_3^1\right\}$ for rule R_1 and $\left\{\mu_2, \beta_1^2, \beta_2^2, \beta_3^2\right\}$ for rule R_2. As the coefficients β_1^1 and β_2^1 determine the contribution of contemporaneous (π_t) and first lagged (π_{t-1}) inflation to the agent's expectations of future inflation, $\hat{E}_t\pi_{t+1}$, it is very important that these beliefs determine a stable AR(2) process. Typically, in a model with only one forecasting rule, as soon as the sum of β_1^1 and β_2^1 exceeds unity, the whole time-varying transmission mechanism of the DSGE model becomes unstable (matrix T_t obtains an eigenvalue outside of the unit circle), which leads to large forecasting errors, invocation of the projection facility[16], and decline in

[16] As many of the forecasting rules that we study are not in the form of autoregressive process, we check for instability at the level of transmission mechanism T rather than beliefs, as in common in the adaptive learning literature. For more details on our approach, consult Slobodyan and Wouters (2012a).

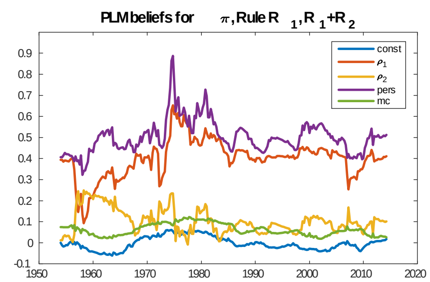

Figure 1: PLM beliefs for inflation, rule R_1, rules R_1 and R_2 in the model.

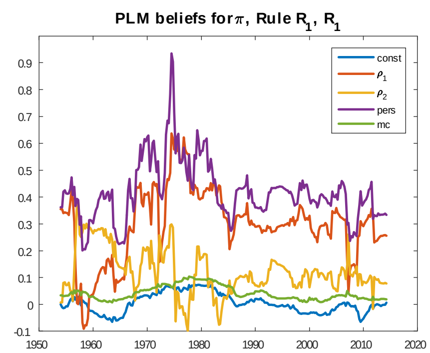

Figure 2: PLM beliefs for inflation, rule R_1, only rule R_1 model.

the likelihood. Therefore, explaining the Great Inflation episodes becomes a challenge for the AL model: on the one hand, its agents are quickly learning that inflation process is approaching a random walk during 70es, on the other, overhitting the unity boundary leads to quick deterioration of the likelihood. The estimation procedure then adjusts the parameters in such a way that the instances of unit boundary hitting are minimized. This struggle is very clearly visible in the Figure 2, where the only rule's persistence is very close to unity around 1975. The model with rules R_1 and R_2, however, has another margin of adjustment: as the rule R_2 involves lagged (π_{t-1}) and twice lagged (π_{t-2}) inflation, breaching the stability boundary isn't as costly in terms of likelihood. In fact, in at least two instances the sum of β_1^2 and β_2^2 is greater than one, and the projection facility is invoked (see Figure 3). However, at the same time the rule R_1 becomes *less* persistent, with the peak persistence falling from 0.935 (Figure 2) to 0.887 (Figure 1). This makes the AL model that combines forecasting rules R_1 and R_2 more robust to disturbances, leading to better posterior probability.

The forecasting rule $R2$ isn't primarily responsible for more flexible transmission mechanism, though. Figure 4, which shows RMSE of inflation forecast

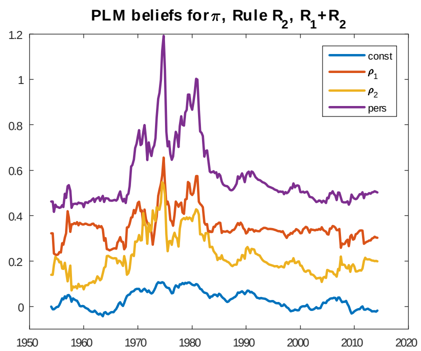

Figure 3: PLM beliefs for inflation, rule R_2, rules R_1 and R_2 in the model.

errors in expanding windows, demonstrates that until the Great Inflation of the 70es, the combined forecasting model (R_1 and R_2) was performing rather well, as the RMSE of the combined forecasting rule (dashed purple line) had better RMSE than the single R_1 rule in a model without R_2 (solid black line). In fact, it is exactly the R_2 rule which was the best until the Great Inflation (solid blue line). After the two high inflation episodes of 70es and 80es, however, this rule became much worse, as its usage of very old information (π_{t-1} and π_{t-2} used to predict π_{t+1}) led to quickly worsening forecasting performance during the Great Inflation and the consequent disinflation. Its performance during this time period is so bad that the rule R_2 never comes close to being the best again.

The above results were obtained when the weights on different rules were fixed. Thus, even when the rule R_2 suffered a sharp deterioration in forecasting performance, its weight in the overall forecasting equation used by adaptively learning agents remained 1/2. Therefore, we turn to the analysis of the results with time varying weights.

First, let's compare the estimation with forecasting rules R_1 and R_2 under fixed and variable weights. Fixed weights estimation is better by about 11 units of the log posterior. This is a rather common situation: AL estimations in Slobodyan and Wouters (2012b) also were better with fixed weights. Figure 5 shows that the expanding window RMSE of rules R_1 and R_2 in case the weights are time-varying behaves very similarly to the case with fixed weights: the model R_2 is the best until early 70es, but quickly deteriorates with the onset of large and volatile inflation. What is remarkable is that RMSE of a combined forecasting rule after late 70es is practically identical to the RMSE of rule R_1, indicating that in that period the agents put very little weight on R_2.

This conjecture is supported by the evolution of forecasting rule weights. The agents begin by quickly pushing the weight for R_1 (solid black line in Figure 6) to one, but then the simpler lagged AR(2) rule R_2 (blue dashed line in Figure 6) starts to be slightly better. Being also smaller, it increases its weight, so

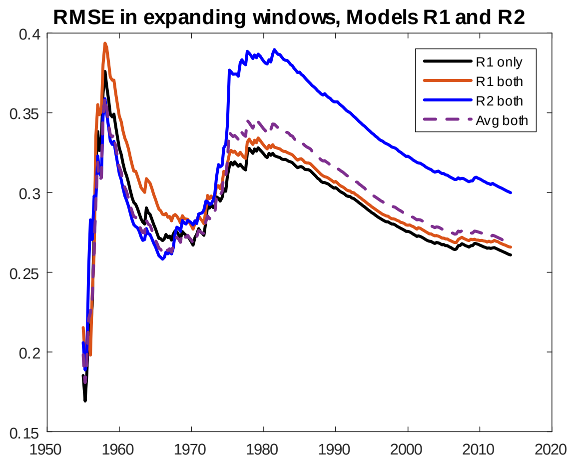

Figure 4: Expanding window Root Mean Squared Errors for forecasting rules R_1 and R_2, only R_1 and $R_1 + R_2$ models. Fixed rule weights.

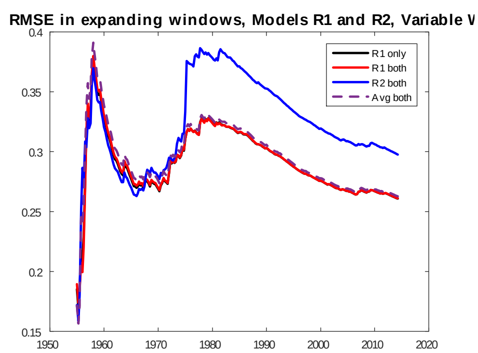

Figure 5: Expanding window Root Mean Squared Errors for forecasting rules R_1 and R_2, only R_1 and $R_1 + R_2$ models. Variable rule weights.

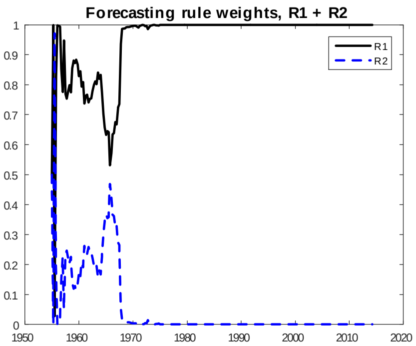

Figure 6: Forecasting rules R_1 and R_2 weights, $R_1 + R_2$ model.

that by late 60es R_1 and R_2 and almost equally important in the agents' minds. However, evolution of the forecasting performance during the 70es eliminates the rule R_2 from the pool of the rules the agents use in their forecasting, as seen in Figure 6.

Estimated models with multiple forecasting rules and variable weights are all rather close to the original baseline in terms of the model fit. It appears that the models involving rule R_3 (the agents are using only the time-varying mean to predict inflation) are closest to the corresponding models with fixed weights which are, however, not the best in terms of log posterior. To evaluate further the reasons for this performance, we look closely at the only model where variable weights estimation delivers better results than fixed weights: the model with forecasting rules R_1 and R_3.

Figures 7 and 8 show the agents' beliefs for forecasting rules R_1 and R_3 in estimation with fixed (Figure 7) and variable (Figure 8) weights. The solid dashed line in both panels is the only beliefs of the R_3 model (the constant), while the rest of the beliefs are the same as in the Figures 1 and 2. It is obvious that the beliefs about the mean inflation in the rule R_3 are varying widely, and more so in the fixed weights estimation (the constant term for the rule R_1 is also rather volatile in this case). The emphasized role of the constant in inflation beliefs when the weights are fixed is accompanied by a significantly lower variability of perceived inflation persistence, which mostly stays around 0.7 and never reaches above 0.85. This range increases to $0.4 \div 0.8$ for the variable weights estimation. Contrast this behavior with the Figure 1, depicting the results for models with R_1 alone and R_1 plus R_2, where the perceived persistence of inflation is generally varying between 0.4 and 0.9. Apparently, the R_1 plus R_2 estimation is striking the right balance in inflation persistence, preventing it from approaching the unit root, but also allowing for sufficient time variation to account for quite and volatile times within the same forecasting rule. The model with the rules R_1 and R_3 is forced to track both quiet and rapidly changing times

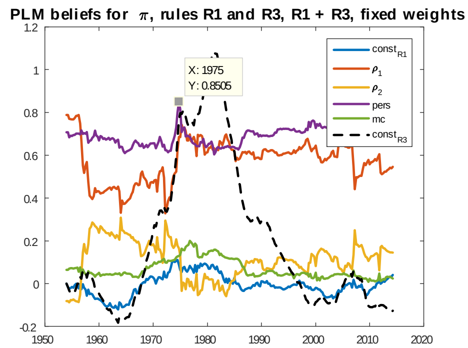

Figure 7: PLM Beliefs for inflation, rules R_1 and R_3, rules R_1 and R_3 in the model. Fixed weights.

mostly with the constant terms, which cannot generate the same time varying volatility that is required to explain the behavior of the data.

Finally, we have to note that Hommes and co-authors use the forecasting rules with fixed beliefs, with the share of agents using different rules changing over time. Our adaptive learning approach implies that the beliefs are time varying, while the weights (analog of the shares of agents using particular rules) could change or stay constant. The results presented in this paper show that it is indeed very important that the agents' beliefs change over time, as the estimation procedure selects parameters values which make them highly variable.[17] In order to have the closest possible comparison with Hommes and co-authors, we now run a series of simulations where forecasting rules R_2, R_3, and R_4 are fixed for the duration of the estimation, with their initial values still given by the REE of the overall model.

This estimation is slightly worse than the baseline with just the R_1 forecasting rule (log posterior -1980.32). This, however, is rather close to the estimation with all four rules and time-varying weights, when the coefficients of the rules are allowed to vary, as well (-1978.81). The pattern of weights (see Figure 9) is very similar to the one that was observed in previous estimations: before 1970, there are several rules which are competing for the agents' usage — R_2 at the very beginning of the estimated sample, and R_3 increasingly important in the 60es, but the quickly disappearing from the set of rules used due to its poor performance. The disappearance occurs despite the fact that the R_3 is very simple, using just an estimate of the inflation mean, which favors it in comparisons that use BIC criterion.

All our results, taken together, suggest that any improvement in the model fit that we get is likely coming from adjusted transmission mechanism of the

[17] The most important parameter governing the time variation of the beliefs is ρ. If it is estimated to be below $0.8 \div 0.9$, the beliefs remain approximately constant over the estimation period.

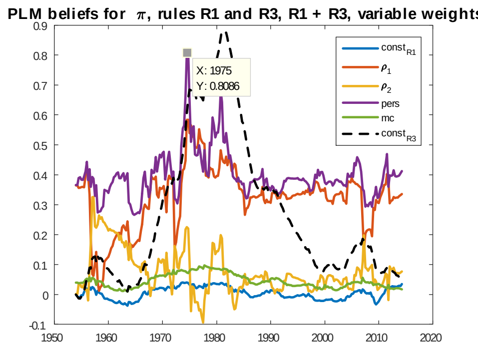

Figure 8: PLM Beliefs for inflation, rules R_1 and R_3, rules R_1 and R_3 in the model. Variable weights.

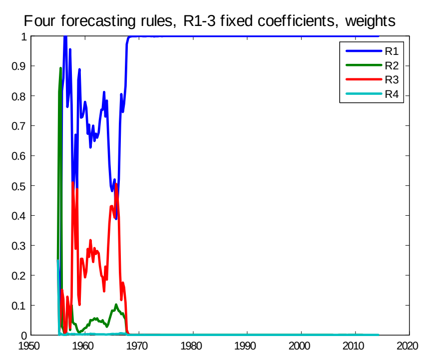

Figure 9: Forecasting rules R_1 through R_4 weights, $R_1 + R_2 + R_3 + R_4$ model. Beliefs for rules R_2, R_3, and R_4 fixed.

model, in particular, the fact that the main forecasting rule, R_1, is less likely to approach the instability boundary in presence of other forecasting rules the agents might use. Actual forecasting performance of the alternatives seems to be comparable to the R_1 only in the relatively tranquil times, before late 60es, but then quickly deteriorates, so that the weight on these alternatives collapses to zero for the rest of the sample. As observed previously by Slobodyan and Wouters (2012b), allowing the agents to select among different rules does not in itself lead to better model fit; it is likely that any improvement comes from modification in the transmission mechanism, in particular the model's ability to better compute time variation in the persistence and volatility of the inflation process.

5 Discussion

As stated previously, this paper is motivated by findings of Hommes and co-authors who identify several simple forecasting rules which experimental subjects seem to be using in the laboratory. A common feature of these experiments (cf., Anufriev and Hommes 2012 or Bao *et al.* 2012) is that there is no simple rule which always dominates an experimental run; moreover, there are many runs in which several rules continue to coexist for the entire duration of the experiment. In our estimations it is always the R_1 model that is left standing after 1970 or so; no other rule survives the Great Inflation period. Still, the presence of other forecasting rules could be beneficial for the overall model fit, as documented in the Table 1.

What explains such a discrepancy in the evolution of a set of forecasting rules that survives the agents' selection process? First, in our estimations there is a difference in the timeliness of the information set that different agents are using: R_1 uses some time t information to forecast the value of π_{t+1}, while only $t-1$ variables are present in the rules R_2 through R_4. We do not consider this to be a vital difference, because our sample includes both quiet (periods where all rules might be similar) and volatile (periods when timeliness is presumably vital) periods. In fact, it is in 50es and 60es, when the inflation was not volatile (but also not persistent), that the other forecasting rules were used with non-zero weight. With the advent of 70es and much higher predictability of inflation,[18] informational advantage of R_1 led to its dominance. Still, long period of the Great Moderation did not lead to resurrection of the other forecasting rules. However, we also note that there were set-ups in the work of Hommes and co-authors, for example in Anufriev and Hommes (2012), when it was time t information that was used to predict $t+1$ variable. Qualitatively, their results stayed similar to the previous work.

In order to test the hypothesis that it is indeed the informational advantage

[18] Our adaptively learning agents perceive that in the 70es the inflation Data Generating Process was approaching a random walk, which makes inflation more predictable than autoregressive process with rather low autocorrelation, the situation observed for most of the remainder of our sample.

that precludes other rules from surviving during the entire sample, we adjust R_1 and let it use only π_{t-1} and π_{t-2} to forecast π_{t+1}. We call this modified rule $\tilde{R}_1$. Somewhat surprisingly, the results are better in the estimation that uses $\tilde{R}_1$ plus (R_2, R_3, R_4) with fixed belief coefficients than in the same estimation with R_1 (-1972.18 vs. -1980.32). The rule R_3, being smallest, is still competing the longest, but by 1970 its weight goes to zero. Thus, the timely information is not an issue.

Another difference between our estimations and Hommes and co-authors work is that we allow the belief coefficients to vary over time. Typically, during the Great Inflation period the agents' beliefs about inflation approach random walk, while outside of this period they are much less persistent. In contrast, all the rules in, for example, Anufriev and Hommes (2012), are equivalent to the random walks.[19] In our estimations the dominant root of all our forecasting rules is less than unity for absolute majority of time periods. It remains an open question whether this difference could be crucial for the fact that only one forecasting rule ever survives in our estimations.[20]

6 Conclusion

In this paper we have estimated a Smets-Wouters type model with financial frictions, developed in Rychalovska et al (2015), with adaptively learning agents. We modeled the agents as choosing among a set of potential forecasting rules, with the share of agents selecting a particular rule being a function of the rule's complexity and past performance. We discovered that our model invariably predicts that the default forecasting rule for inflation, utilizing a constant, two lags of inflation, and marginal cost variable, is the best one over the whole estimated sample (1954Q1 to 2014Q2). Its advantage comes despite the fact that this is one of the most complex rules used, which is penalized for its complexity through using BIC as a selection criterion.

We discussed the reasons for such uniformity, which is contrasting with the results of learning-to-forecast experiments by Hommes and co-authors who typically find that a wider set of forecasting rules survives, and that there is no a single rule that always dominates an experimental session. We speculate that the

[19] We list their rules and provide our closest analogues below:

$$
\begin{aligned}
(R_3) \qquad & p^e_{t+1} = p^e_t + 0.65 \cdot (p_t - p^e_t)\,, \\
(R_2) \qquad & p^e_{t+1} = p_t + 0.40 \cdot (p_t - p_{t-1})\,, \\
(R_2) \qquad & p^e_{t+1} = p_t + 1.30 \cdot (p_t - p_{t-1})\,, \\
(R_4) \qquad & p^e_{t+1} = 0.5 \cdot (p^{av}_t + p_t) + (p_t - p_{t-1})\,.
\end{aligned}
$$

[20] We note that the agents in Hommes and co-authors experiments are predicting price level rather than inflation. Even though their underlying model predicts a finite price in equilibrium, it is unclear whether the forecasting rules that they fit to the results could enforce unbounded price dynamics.

However, experimental subjects, participating in the experiment, could consider price level as a variable that is not necessarily bounded. It is unlikely that the actual economic agents, who according to our model had to predict inflation, could believe in its non-stationarity, even during the times of the Great Inflation.

difference could be due to, first, the fact that the data does not support most of their rules being dominant,[21] and second, that it could be determined by experimental agents predicting price level, which could be thought of a non-stationary, while the inflation beliefs were typically predicting stable (stationary) dynamics of this variable. We plan to continue working on the issue to further understand the reason for this divergence.

Further extensions of this project will include incorporating a variety of forecasting rules for *all* forward-looking variables, in contrast to just inflation rules which were studied in this paper. We also plan to study whether shortening the effective window that the agents use to measure forecasting performance of a particular rule has an effect on the time profile of the rules' usage.

References:

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," Econometrics Reviews, Taylor and Francis Journals, vol. 26(2-4): 113-172.

Anufriev, M. and Hommes C. H. (2012), "Evolutionary selection of individual expectations and aggregate outcomes," American Economic Journal: Microeconomics, 4(4), 35-64.

Bao, T., Hommes, C. H., Sonnemans, J., and J. Tuinstra. (2012), "Individual expectations, limited rationality and aggregate outcomes," Journal of Economic Dynamics & Control, 36(8), 1101-1120.

Bao, T., Hommes, C. H., and T. A. Makarewitz (2016), "Bubble formation and (in)efficient markets in learning-to-forecast and -optimize experiments," Economic Journal, forthcoming.

Bernanke, B. and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," American Economic Review, Vol. 79 (1): 14-31

Bernanke, B., M. Gertler, and S. Gilchrist (1999): "The financial accelerator in a quantitative business cycle framework" in J Taylor and M Woodford (eds) Handbook of Macroeconomics.Volume 1C, North Holland.

Brzoza-Brzezina, M. and M. Kolasa (2013): "Bayesian Evaluation of DSGE Models with Financial Frictions," Journal of Money, Credit and Banking, Blackwell Publishing, vol. 45(8): 1451-1476, December.

Carlstrom, C., and T. Fuerst (1997): "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis," American Economic Review, vol. 87(5): 893-910, December.

Cespedes, L., R. Chang and A. Velasco (2004): "Balance sheets and exchange rate policy," The American Economic Review, Vol. 94(4), September.

Christiano, L.J. , R. Motto and M. Rostagno (2003): "The great depression and the Friedman–Schwartz hypothesis," Journal of Money, Credit and Banking,

[21] Inflation dynamics over the sample exhibits convergence characteristics after the Great Inflation episode was over. Hommes and others typically find that if the price dynamics in the experimental run has converged, it is a particular rule (simple adaptive expectations) that tends to dominate. Therefore, one does not expect other than the simple adaptive expectations rules to prevail in convergent periods, which prevailed after the mid-80es.

vol. 35: 1119–1197.

Christiano, L.J., M. Eichenbaum and C.L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," Journal of Political Economy, Vol. 113(1): 1–45.

Christiano, L.J., R. Motto, and M. Rostango (2010): "Financial factors in economic fluctuations," ECB Working Paper 1192.

Christensen, I., and A. Dib (2008): "The financial accelerator in an estimated New Keynesian model," Review of Economic Dynamics 11: 155-178.

De Graeve, F., (2008): "The external finance premium and the macroeconomy: U.S. post-WWII evidence," Journal of Economic Dynamics and Control vol. 32.

Evans, G., and S. Honkapohja (2001): "Learning and Expectations in Macroeconomics," Princeton University Press, Princeton, NJ.

Gilchrist, S., A. Ortiz, and E. Zakrajsek(2009): "Credit risk and the macroeconomy: Evidence from an estimated DSGE model," mimeo.

Haario, H., E.Saksman, and J.Tamminen (2001): "An Adaptive Metropolis Algorithm," Bernoulli, Vol.7(2): 223-242.

Hommes, C. H. (2011): "The heterogeneous expectations hypothesis: some evidence from the lab," Journal of Economic Dynamics and Control 35, 1-24.

Kiyotaki, N., and J.Moore (1997): "Credit cycles," Journal of Political Economy, vol. 105: 211-248.

Merola R., 2013. "The role of financial frictions during the crisis: An estimated DSGE model," Working Paper Research 249, National Bank of Belgium.

Milani, F., (2006): "A Bayesian DSGE Model with Infinite-Horizon Learning: Do 'Mechanical' Sources of Persistence Become Superfluous?" International Journal of Central Banking, Iss. 6.

Milani, F., (2007): "Expectations, Learning and Macroeconomic Persistence", Journal of Monetary Economics, vol. 54: 2065–2082.

Milani, F., (2008): "Learning, Monetary Policy rules, and Macroeconomic Stability," Journal of Economic Dynamics and Control, 32(10): 3148-3165.

Iacoviello, M., (2005): "House prices, borrowing constraints and monetary policy in the business cycle," American Economic Review, 95(3): 739-764.

Iacoviello, M., and S.Neri (2008): "Housing market spillovers: evidence from the estimated DSGE model," Bank of Italy, WP 659.

Orphanides, A., and J. Williams (2007): "Inflation targeting under imperfect knowledge," Economic Review, Federal Reserve Bank of San Francisco, pages 1-23.

Rychalovska, Y., and S. Slobodyan (2010): "Bayesian estimation of DSGE models for the Euro Area: robustness issues," mimeo.

Rychalovska, Y., S. Slobodyan, and R. Wouters (2015): "Financial Frictions, Expectations and Business Cycle: Evidence from an Estimated DSGE Model," mimeo.

Sargent, T.J., N. Williams, and T. Zha (2006): "Shock and Government Beliefs: The Rise and Fall of American Inflation," American Economic Review, vol. 96(4): 1193-1224.

Smets, F., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," Journal of the European Economic Association, vol. 1(5): 1123-75.

Smets, F., and R. Wouters (2004): "Forecasting with a Bayesian DSGE Model: An Application to the Euro Area," Journal of Common Market Studies, vol. 42(4): 841-67.

Smets, F., and R. Wouters (2007): "Shocks and frictions in U.S. business cycle: A Bayesian DSGE approach," American Economic Review, vol. 97(3).

Slobodyan,S., and R. Wouters (2012a): "Learning in an estimated medium-scale DSGE model," Journal of Economic Dynamics and Control vol.36: 20-46.

Slobodyan,S., and R. Wouters (2012b): "Learning in a medium-scale DSGE model with expectations based on small forecasting models," American Economic Journal: Macroeconomics vol 4(2): 1-40.

Slobodyan,S., and R. Wouters (2017): "Survey expectations of inflation and learning," mimeo.

Soto, C., J.P. Medina, and R. Caputo (2010): "The financial accelerator under learning and the role of monetary policy," Central Bank of Chile, Working Paper 590.

7 Model appendix

7.1 Households and Labor Markets

Household j chooses consumption, hours worked and savings so as to maximize a utility function, non-separable[22] in two arguments – a CES basket of consumption-good varieties and labor services:

$$E_t \sum_{s=0}^{\infty} \beta^s \left[\frac{1}{1-\sigma_c} (C_{t+s}(j) - \eta C_{t+s-1})^{1-\sigma_c} \right] \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_{t+s}(j)^{1+\sigma_l} \right), \quad (13)$$

where σ_c and σ_l are preference parameters and η is an external habit-formation parameter, which introduces the dependence of the household consumption on the lagged aggregate consumption. Households can save by depositing funds in the bank and by buying government bonds. These assets (denoted, in total, as AT) are perfect substitutes and earn the same riskless nominal interest rate R^n. Households also obtain dividends from owing intermediate and capital goods producers as well as from labor unions. Therefore, the budget constraint of the representative household takes the form:

$$C_{t+s}(j) + \frac{AT_{t+s}(j)}{\varepsilon_t^b R_{t+s}^n P_{t+s}} - T_{t+s} = \frac{W_{t+s}(j) L_{t+s}(j)}{P_{t+s}} + \frac{AT_{t+s-1}(j)}{P_{t+s}} + \frac{Div_{t+s}}{P_{t+s}}, \quad (14)$$

where c_t is the exogenous premium on the bonds' return, W_{t+s}^h is the nominal wage, T_{t+s} are lump-sum taxes or subsidies and Div_{t+s} are dividend payments.

[22] A sensitivity check demonstrated that the use of the separable (in consumption and labor) form of the utility function, employed in Smets and Wouters (2003), does not significantly affect the estimation results or the conclusions of the paper.

The first-order conditions with respect to consumption and assets result in the Euler equation, which after model detrending and log-linearization takes the following form:

$$\widehat{c}_t = \frac{1}{(1+(\eta/\gamma))}E_t\left[\widehat{c}_{t+1}\right] + \frac{(\eta/\gamma)}{(1+(\eta/\gamma))}\widehat{c}_{t-1} \tag{15}$$
$$- \frac{(1-\eta/\gamma)}{\sigma_c(1+(\eta/\gamma))}(\widehat{b}_t + \widehat{R}_t^n - E_t[\widehat{\pi}_{t+1}]) - \frac{(\sigma_c-1)(w_*^h L/c_*)}{\sigma_c(1+(\eta/\gamma))}(E_t\left[\widehat{L}_{t+1}\right] - \widehat{L}_t).$$

The backward-looking term arises in the consumption equation due to the assumptions of external habit formation captured by the parameter η. Therefore, current consumption ($\widehat{c}_t$) depends on a weighted average of past and expected future consumption. The consumption process is also affected by the expected growth in hours worked ($E_t\left[\widehat{L}_{t+1}\right] - \widehat{L}_t$) (due to the non-separable in consumption and labor form of the utility function), the ex–ante real interest rate ($\widehat{R}_t^n - E_t[\widehat{\pi}_{t+1}]$) and a disturbance term $\widehat{b}_t$. γ is the deterministic trend, which arises as a result of model detrending[23]. $\widehat{b}_t$ is assumed to follow a first–order autoregressive process with an iid–normal error term: $\widehat{b}_t = \rho_b \widehat{b}_{t-1} + \epsilon_t^b$. Variables with stars denote the steady-state values.

As in Smets and Wouters (2007), labor markets consist of labor unions, who allocate and differentiate labor supplied by households, and labor packers, who buy labor from the unions, package it into a Kimball (1995) composite aggregator L_t that is resold to intermediate goods producers. Unions have market power over labor services and set wages that are subject to nominal rigidities along the lines of Calvo (1983). Every period only a $(1-\xi_w)$ fraction of intermediate labor unions can readjust wages. The chosen wage rate set by the union maximizes the stream of future (discounted) wage income for all the time periods when the union is stuck with that wage in the future. The first-order conditions to problems (13) and (14) with respect to hours worked combined with the solution to the profit-maximization problem of the intermediate labor union and the law of motion of the aggregate wage result in the following wage equation:

$$\widehat{w}_t = \frac{1}{(1+\overline{\beta}\gamma)}(\widehat{w}_{t-1} + \overline{\beta}\gamma E_t\left[\widehat{w}_{t+1}\right] - (1+\overline{\beta}\gamma\iota_w)\widehat{\pi}_t + \iota_w\widehat{\pi}_{t-1} + \overline{\beta}\gamma E_t\left[\widehat{\pi}_{t+1}\right] \tag{16}$$
$$+ \frac{(1-\xi_w\overline{\beta}\gamma)(1-\xi_w)}{\xi_w((\phi_w-1)\varepsilon_w+1)}\left[\frac{1}{1-\eta/\gamma}\widehat{c}_t - \frac{\eta/\gamma}{1-\eta/\gamma}\widehat{c}_{t-1} + \sigma_l\widehat{L}_t - \widehat{w}_t\right] + \widehat{\lambda_{w,t}},$$

where $\overline{\beta} = \beta/\gamma^{\sigma_c}$ and β is a discount factor applied to households. Due to nominal wage stickiness and the partial indexation of wages to inflation, real wages adjust only gradually to the desired wage mark–up. ξ_w is a wage stickiness parameter. Parameter ι_w measures the degree of indexation. If wages are perfectly flexible ($\xi_w = 0$), the real wage is a constant mark–up over the marginal

[23] Detrended real variables are obtained by dividing the nominal variables by a deterministic trend: $c_t = C_t/\gamma^t$, $w_t = W_t/(\gamma^t P_t)$ etc.

rate of substitution between consumption and leisure. When wage indexation is zero (ι_w), real wages do not depend on lagged inflation. In addition to wage stickiness, the speed of adjustment to the desired mark-up depends on the demand elasticity for labor, which is a function of the steady-state labor market mark-up ($\phi_w - 1$) and the curvature of the Kimball labor-market aggregator ε_w. The wage–mark up disturbance ($\widehat{\lambda}_{w,t}$) is assumed to follow an ARMA (1,1) process with an iid–normal error term: $\widehat{\lambda}_{w,t} = \rho_w \widehat{\lambda}_{w,t-1} - \mu_w \epsilon_{w,t-1} + \epsilon_t^w$.

7.2 Production sector: Firms

The production sector consists of final- and intermediate -goods producers. Final -goods producers buy intermediate goods $Y_t(i)$, aggregate them into a composite final good Y_t and resell to consumers in a perfectly competitive market. The solution to the profit-maximization problem of these firms is standard and determines the demand function for intermediate inputs $Y_t(i)$. Intermediate -goods producers, who operate under monopolistic competition, rent capital from entrepreneurs at the rate R_t^k, hire labor from labor packers and use a typical Cobb-Douglas production function augmented with fixed costs:

$$Y_t(i) = \varepsilon_t^a K_t^S(i)^\alpha \left[\gamma^t L_t(i)\right]^{1-\alpha} - \gamma^t \Phi, \tag{17}$$

where $K_t^S(i)$ is capital services used in production, $L_t(i)$ is aggregate labor input, α is the share of capital in production and Φ is a fixed cost. γ^t respresents the labor-augmenting deterministic growth rate in the economy and ε_t^a is total factor productivity. The log-linearized aggregate supply equation (17) takes the form:

$$\widehat{y}_t \; = \Phi(\; \alpha(\widehat{k^S}_t) \; + \; (1-\alpha)\widehat{L}_t + \widehat{A}_t), \tag{18}$$

where the total factor productivity ($\widehat{A}_t$) is assumed to follow a first-order autoregressive process: $\widehat{A}_t \; = \rho_a \widehat{A}_{t-1} + \epsilon_t^a$. The solution to the cost-minimization problem yields the conditions that determine the labor demand function in the following log-linear form:

$$\widehat{L}_t = \widehat{k^S}_t \; - \; \widehat{w}_t \; + \; \widehat{r}_t^k. \tag{19}$$

Equation (19) implies that the rental rate of capital is negatively related to the capital-labor ratio and positively to the real wage (both with unitary elasticity). The marginal cost is the same for all firms and represented by the following relation:

$$\widehat{mc}_t \; = \; (1-\alpha) \; \widehat{w}_t + \alpha \; \widehat{r}_t^k - \widehat{A}_t. \tag{20}$$

Similar to wages, in each period, only a fraction of firms $(1-\xi_p)$ can re-optimize prices. In the environment of price rigidities, the optimal price will maximize the expected discounted stream of future profits for the firm for all states of nature when the firm cannot reset the price optimally. Thus, the current inflation rate will depend on current *and* future expected marginal costs. Nonreoptimized prices are partially index-linked to past inflation, which gives rise to the backward-looking term in the inflation equation. Profit maximization by price–setting intermediate firms gives rise to the following New-Keynesian Phillips curve:

$$\widehat{\pi}_t = \frac{1}{(1+\overline{\beta}\gamma\iota_p)}(\iota_p\widehat{\pi}_{t-1}+\overline{\beta}\gamma E_t\left[\widehat{\pi}_{t+1}\right]+\frac{1}{((\phi_p-1)\varepsilon_p+1)}\frac{(1-\xi_p\overline{\beta}\gamma)(1-\xi_p)}{\xi_p}(\widehat{mc}_t))+\widehat{\lambda_{p,t}}, \tag{21}$$

where ι_p denotes the indexation coefficient. The inflation equation demonstrates that the speed of adjustment to the desired mark-up depends on the degree of price stickiness ξ_p, the curvature of the Kimball goods market aggregator ε_p and the steady state mark-up $(\phi_p - 1)$. The price mark–up disturbance $(\widehat{\lambda}_{p,t})$ is assumed to follow an ARMA(1,1) process: $\widehat{\lambda}_{p,t} = \rho_p\widehat{\lambda}_{p,t-1} - \mu_p\epsilon_{p,t-1} + \epsilon_t^p$, where ϵ_t^p is an iid–Normal price mark–up shock.

7.3 Monetary policy and equilibrium

The model is completed by adding the following empirical monetary policy reaction function:

$$\begin{aligned}\hat{R}_t^n \;=\; \rho_R\hat{R}_{t-1}^n \;+\; (1-\rho_R)(r_\pi\widehat{\pi}_t \;+\; r_y\widehat{ygap}_t) \\ +\, r_{\Delta y}(\widehat{ygap}_t - \widehat{ygap}_{t-1}) \;+ r_t.\end{aligned} \tag{22}$$

The monetary authority follows a generalized Taylor rule responding to inflation and the output gap terms (current and lagged). The latter is defined as the difference between actual and potential output. The output gap is approximated by $\widehat{ygap}_t = \widehat{y}_t - \hat{A}_t$. The parameter ρ_R captures the degree of interest rate smoothing, and the monetary policy shock (r_t) follows a first-order autoregressive process with an iid–Normal error term: $\widehat{r}_t \;= \rho_r\widehat{r}_{t-1} + \epsilon_t^r$.

The log-linear representation of the resource constraint is given by

$$\widehat{y}_t \;=\; \frac{(\overline{R}^K - 1 + \tau)k_*}{y_*}\widehat{u}_t + \widehat{\mu}_t^{bank} + \;\frac{c_*}{y_*}\widehat{c}_t \;+\; \frac{i_*}{y_*}\widehat{i}_t \;+\widehat{g}_t \;, \tag{23}$$

where $\widehat{\mu}_t^{bank} = (k_*/y_*)(\overline{R}^K - \overline{R})(1-\overline{N}/\overline{K})(\hat{R}_t^K + \widehat{Q}_{t-1} + \widehat{k}_t)$ measures the bank monitoring cost, and $\widehat{g}_t$ is exogenous government spending, following an AR(1) process with an iid–normal error term, also affected by the productivity shock as in Smets and Wouters (2007), $\widehat{g}_t \;= \rho_g\widehat{g}_{t-1} \;+ \rho_{ga}\epsilon_t^a + \epsilon_t^g$.