Working Paper Series 792
(ISSN 2788-0443)

Dynamic Sparse Restricted Perceptions Equilibria

Volha Audzei
Sergey Slobodyan

CERGE-EI
Prague, October 2024

Dynamic Sparse Restricted Perceptions Equilibria

Volha Audzei and Sergey Slobodyan*

September 18, 2024

Abstract

This paper studies convergence properties, including local and global strong E-stability, of the rational expectations equilibrium under non-smooth learning dynamics. In a simple New Keynesian model, we consider two types of informational constraints operating jointly - adaptive learning and sparse rationality. For different initial beliefs, we study if the convergence to the minimum state variable rational expectations equilibrium (MSV REE) occurs over time for positive costs of attention. We find that for any initial beliefs the agents' forecasting rule converges either to the MSV REE equilibrium, or, for large attention costs, to a rule that disregards all variables but the constant. Stricter monetary policy slightly favors the constant only rule. Mis-specified forecasting rule that uses variable not present in the MSV REE does not survive this learning algorithm. Theory of non-smooth differential equations is applied to study the dynamics of our learning algorithm.

Keywords: Bounded rationality, Expectations, Learning, Monetary policy.

JEL: D84, E31, E37, E52.

*Audzei: Czech National Bank, Na Prikope 864/28, Prague, Czech Republic , vaudzei@cerge-ei.cz. Slobodyan: CERGE-EI, Politickych veznu 7, 111 21, Prague, Czech Republic, sergey.slobodyan@cerge-ei.cz. The views expressed in this paper are those of the authors and not necessarily those of the Czech National Bank. The authors thank Seppo Honkapohja and participants of the SCE 2023 conference for valuable comments.

1 Introduction

1.1 Sparse Adaptive Learning

Modelling agents' expectation formation in self-referential systems has been extensively debated in the literature. The debate has received a different angle when increased computer power and data availability simplified the tasks of dimensionality reduction and variable selection.[1] In a dynamic self-referential system, the choice of variables in the forecasting models affects the system itself. A vast Adaptive Learning (AL) literature[2] has studied the conditions under which such a system of forecasting models converge to a true underlying model. The convergence has been shown to depend on the stability of ordinary differential equations characterizing the underlying model, and the concept has been named Expectational-, or E-stability.

In this paper, we contribute to the AL literature by addressing the global stability of the Rational Expectations Equilibrium (REE) taking into account a non-smooth learning algorithm, where the agents start with a mis-specified forecasting rule and can switch model specification during learning. We study the dynamics of a simple New Keynesian model with boundedly rational agents, who operate under a combination of two types of information constraints: Recursive Least Squares (RLS) learning, with the agents updating their beliefs about the coefficients in their forecasting rules, and Sparse Rationality, which imposes costs on the attention weights of different variables in the forecasting rule, thus selecting variables to be used in the rule. We call this algorithm Sparse Adaptive Learning (Sparse AL). We are interested in the dynamics of the model, which starts with initial beliefs that are significantly different from those consistent with the Minimum State Variable Rational Expectations (MSV REE) equilibrium developed by McCallum (1983, 2003). In particular, motivated by Audzei and Slobodyan (2022), we let our agents start arbitrarily close to the ”wrong” Restrictive Expectations Equilibrium (RPE), in which case the initial forecasting rule includes a variable that is absent from the MSV REE solution. If convergence to the MSV REE solution occurs for arbitrary initial beliefs, one could speak of global - as opposed to local, or asymptotic - strong Expectational Stability of the MSV REE with respect to the RPE forecasting rules under a sparsity information constraint. We also derive the extended version of the E-stability condition for our Sparse Adaptive Learning set-up, and investigate consequences of the relative speed of adjustment of the beliefs and of the attention weights on the convergence properties of their joint dynamics. We further consider the role of monetary policy in the model dynamics and the equilibrium selection.

The Sparse Rationality approach formulated in Gabaix (2014) is a type of penalized regression,

[1] Examples include principal components analysis: Stock and Watson (2002), dynamic factor models: Banbura et al. (2013), Stock and Watson (2016)), and active spaces or penalized regressions: see Hansen and Liao (2019), Korobilis (2013), and Nazemi and Fabozzi (2018). A systematic way of using many variables in econometric and forecasting models is represented by many variants of penalized regressions, including Ridge, Lasso, elastic nets, etc,: see Gefang (2014), Tibshirani (1996), De Mol et al. (2008), or Yuan and Lin (2007). See also Andrle and Bruha (2023) for sparse Kalman filter estimation.

[2] For an overview of Adaptive Learning literature in macroeconomics see Evans and Honkapohja (2001).

specifically a non-negative garrote. A penalized regression minimizes loss function that consists of the sum of squared prediction errors plus a penalty term. In garrote the penalty is imposed on the sum of absolute values of *attention weights* on different variables. Solving the problem of minimizing the loss function leads to derivation of optimal attention weights, some of which could be zero, thus inducing sparsity. The penalty could then be interpreted as the cost of attention that comes from cognitive limitations or cost of data collection.

We model our agents as engaged in Sparse AL, constantly updating their estimates of the forecasting rules (beliefs and attention weights) using an expanding window of the data and Recursive Least Squares (RLS).[3] The estimates are then used to form expectations of the future values of macroeconomic variables and the agents' actions which affect future realizations of data. The process of expectations formation then becomes self-referential. With Sparse AL, the agents are not only updating the regression estimates, but are also constantly making a decision on the amount of attention to be paid to different variables, thus adding another mechanism into the usual self-referential feedback loop studied in the adaptive learning literature.

We find that MSV REE is generally strongly E-stable under Sparse AL. However, for larger values of the attention cost or more aggressive monetary policy, the forecasting rules consisting of the constant only could become stable. Thus, the monetary policy could affect what variables the agents pay attention to, in line with the findings of Audzei and Slobodyan (2022). In contrast to that paper, however, we find that the area in the parameter space consistent with the MSV REE forecasting rule is significantly larger once we allow for learning dynamics under attention costs constraints. We also find that allowing for learning under Sparse Rationality asymptotically eliminates the RPE, which turns out to be very fragile to the introduction of extra variables.

1.2 Literature Review

This paper is related to a large strand of literature on AL and its interaction with monetary policy. A summary of the AL approach is provided in Evans and Honkapohja (2001). Seminal contributions related to interaction with monetary policy include Orphanides and Williams (2007), who studied the robustness of monetary policy rules when agents are learning. The monetary policy analysis in our paper is related more to the studies that address how monetary policy affects the learnability and stability of the equilibria under the learning process: see Mele et al. (2020), Bullard and Mitra (2002), Slobodyan et al. (2016) and Gibbs (2017). In these studies, the learnability and stability of a desired equilibrium is viewed as additional desiderata for a monetary policy rule. We further relate to the studies on survival of mis-specified equilibra in self-referential systems. Evans et al. (2012) showed that the convergence to a mis-specified equilibrium occurs when the feedback parameter on the expectations is strong. A similar conclusion was obtained by Adam (2005), Hommes and Zhu

[3] A variety of other variants is possible, such as Constant Gain learning, Kalman Filter learning, etc.

(2014), Hommes (2014) Branch et al. (2022) and Hajdini (2022), but under different formulations of mis-specification and learning process. In Audzei and Slobodyan (2022) we have shown that the stronger monetary policy response to inflation, which is inversely related to the expectational feedback parameter, makes the survival of a mis-specified equilibrium less likely. We contribute to this literature by adding the variable selection and sparsity considerations to Adaptive Learning, and by studying the global (Sparse) E-stability of the resulting dynamic system.

Our formulation of the initial mis-specified equilibrium is inspired by empirical and theoretical literature on RPE. Studies have demonstrated that models with the forecasters using simple prediction rules for inflation outperform those with the complicated rules in survey and experimental settings: see Branch and Evans (2006), Adam (2007), Hommes (2014), and Pfajfar and Žakelj (2014). Related to the model behavior at the effective lower bound, Ascari et al. (2023) show that combining the RPE and bounded rationality helps to restore the uniqueness of an equilibrium. In the context of estimated New Keynesian models, Slobodyan and Wouters (2012a and 2012b), Audzei (2023), Ormeno and Molnar (2015) and Vázquez and Aguilar (2021), showed that assuming agents use very simple forecasting rules leads to superior model fit in estimated DSGE models under adaptive learning than REE.

In our framework agents' dynamic decisions on including or excluding the variables from the forecasting rules could introduce discontinuity into the model dynamics when the set of included variables changes. This possible discontinuity forces us to rely on the theory of non-smooth differential equations, see Filippov (1988) and Jeffrey (2019). In addition to the standard convergence of the learning dynamics, *sliding* dynamics along the boundary where the agents are indifferent between two different forecasting rules could be observed. The appearance of sliding affects convergence properties of different equilibria.

The paper is structured as follows. We start by describing the model setup and the existence and stability of MSV and an initially mis-specified RPE under AL with expert advice in Section 2. We continue with formulating the sparsity problem of agents choosing variables' attention weights in Section 3. We further study the dynamics of the AL under sparsity, in particular the stability and learnability of different equilibria in Section 4, where we also address the global E-stability of MSV REE using the theory of non-smooth differential equations. We demonstrate analytically how the non-smooth dynamics allows us to compare the model dynamics under fast and slow updating of underlying variable selection in Section 5. Section 6 concludes.

2 The Model

Our model is a standard New Keynesian (NK) model in which consumer utility possesses external habit persistence and a central bank reacts to the deviation of expected inflation from the zero inflation target. The model has been studied extensively; therefore we present below the key equations and leave the detailed derivations to the Appendix A. The model is a three equations NK model, with the following equations:

$$y_t = -\frac{1-h}{(1+h)\sigma}(i_t - E_t\pi_{t+1}) + \frac{1}{1+h}E_t y_{t+1} + \frac{h}{1+h}y_{t-1} + g_t, \tag{1}$$

$$\pi_t = \beta E_t \pi_{t+1} + \omega y_t + u_t, \tag{2}$$

$$i_t = \phi_\pi E_t \pi_{t+1}. \tag{3}$$

Here π is inflation and y the output gap, while the shocks u and g are both *i.i.d.* zero mean random variables with finite variances. The first equation is an investment-savings curve, the second is a new Keynesian Phillips curve, and the last is the central bank reaction function.

We plug the central bank's policy rule (3) into (1)-(2) and rearrange to express the dynamics of inflation and output as a function of their lagged and expected values and shock realizations:

$$y_t = -\frac{1-h}{(1+h)\sigma}(\phi_\pi - 1)E_t\pi_{t+1} + \frac{1}{1+h}E_t y_{t+1} + \frac{h}{1+h}y_{t-1} + \frac{1-h}{(1+h)\sigma}g_t, \tag{4}$$

$$\pi_t = \left(\beta - \frac{\omega(1-h)(\phi_\pi - 1)}{(1+h)\sigma}\right)E_t\pi_{t+1} + \frac{\omega}{1+h}E_t y_{t+1} + \frac{\omega h}{1+h}y_{t-1} + \frac{\omega(1-h)}{(1+h)\sigma}g_t + u_t. \tag{5}$$

To reduce the dimensionality of the problem we assume that agents are only uncertain about inflation dynamics. The forecast of the output gap is obtained in the spirit of Molnar (2007) from outside experts, who posses full knowledge of the system in (4)-(5) and agents' beliefs as described below.

2.1 Expert Forecasts of Output Gap

Following Molnar (2007), we assume that the agents have access to the expert advice on output gap forecasts. These experts are fully aware of the structure of the model and underlying parameters, and make forecasts *given* agents' inflation beliefs.

In Appendix B we characterize the dynamics of output and inflation as a function of expectations of the output gap, given the agents' beliefs about inflation. As experts are fully rational, we solve for their output gap forecasts using the method of undetermined coefficients. As a result, the forecast of the output gap can be written as

$$E_t y_{t+1} = \tilde{\gamma}_y y_{t-1} + \tilde{\gamma}_\pi \pi_{t-1} + \tilde{\gamma}_u u_t + \tilde{\gamma}_g g_t, \tag{6}$$

where the coefficients are functions of the agents' inflation beliefs defined in (B37)-(B40) and (B41)-(B44).

With these output gap expectations, we can re-write the model in (4)-(5) as follows:

$$y_t = -\frac{1-h}{(1+h)\sigma}(\phi_\pi - 1)E_t\pi_{t+1} + \frac{\tilde{\gamma}_\pi}{1+h}\pi_{t-1} + \frac{h+\tilde{\gamma}_y}{1+h}y_{t-1} + \frac{1-h+\sigma\tilde{\gamma}_g}{(1+h)\sigma}g_t + \frac{\tilde{\gamma}_u}{1+h}u_t, \tag{7}$$

$$\pi_t = \left(\beta - \frac{\omega(1-h)(\phi_\pi - 1)}{(1+h)\sigma}\right)E_t\pi_{t+1} + \tag{8}$$

$$\frac{\tilde{\gamma}_\pi\omega}{1+h}\pi_{t-1} + \frac{\omega(h+\tilde{\gamma}_y)}{1+h}y_{t-1} + \frac{\omega((1-h)+\sigma\tilde{\gamma}_g)}{(1+h)\sigma}g_t + (\frac{\omega\tilde{\gamma}_u}{1+h}+1)u_t.$$

The agents' forecasting rules in this paper could be consistent with the MSV REE and only use the output gap, be based on the Restricted Perceptions Equilibrium (RPE) and thus include only inflation, or they could be formed as some linear combination of these two beliefs. We nowproceed to discuss the properties of the relevant equilibria.

2.2 Minimal State Variable Solution

Having defined the forecast of the output gap we begin our analysis by defining and studying the properties of MSV REE solution for inflation. At MSV, the agents' beliefs about the inflation - their Perceived Law of Motion (PLM), contain only the state variable: output gap and observed shocks. Such beliefs are given by:

$$\pi_t = c^y_\pi y_{t-1} + \gamma^y_\pi g_t + u_t, \tag{9}$$

$$y_t = c^y_y y_{t-1} + \gamma^y_y g_t, \tag{10}$$

with the coefficients derived in Appendix C.

When the agents are using the MSV functional form to formulate their PLM[4] and then use the RLS learning algorithm to learn the coefficients in their PLM, the dynamics of their currently assumed coefficients (beliefs) is asymptotically governed by the approximating ordinary differential equation (ODE). Asymptotic stability of a stationary point of this approximating ODE, or E-stability, is related to the convergence of the agents' beliefs to their MSV REE values: see Evans and Honkapohja (2001) and further elaboration in the Appendix C. We show in Appendix C.1 that the MSV REE is weakly E-stable as long as the Taylor principle is satisfied, $\phi_\pi > 1$.[5]

In this paper we are mostly interested in a form of global strong E-stability concept of the MSV solution. The weak E-stability just described guarantees only that if the agents' forecasting rule has the same functional form as the MSV solution, and the values of coefficients they believe in are contained in some small neighborhood of the MSV values initially, then under appropriate assumptions the RLS

[4]We are using mixed dating. When the agents form expectations and thus assume that only y_{t-1} matters for forming forecasts of future inflation and output gap, given that the shocks are assumed to be *i.i.d.*

[5]For a very persistent output dynamics, as shown in the Appendix C.1, the condition on ϕ_π is stricter than the usual Taylor principle.

learning will converge to MSV REE with a probability approaching 1, see Evans and Honkapohja (2001). This result leaves two unanswered questions. First, what happens if the initial forecasting rule contains more variables than the MSV set? Will the agents asymptotically learn that the value of extraneous coefficients in the PLM are zero? In other words, does the *strong* E-stability obtain? And second, what happens if the initial beliefs are far away from the MSV ones: will we still observe convergence? In other words, is the weak or strong E-stability not just *local (asymptotic)*, but *global*?

Contrary to the weak E-stability, a strong E-stability concept is not unique. It is defined only with respect to the specific mis-specification of the PLM from which the learning is assumed to begin. The Appendix C.1 derives conditions under which the strong E-stability obtains when the agents allow another endogenous variable, π, to be present in their PLM. If the sufficient condition for the weak E-stability is satisfied, strong E-stability obtains as well. With strong E-stability, the beliefs will still converge to the correct MSV REE beliefs and the agents will learn that the inflation doesn't belong in their PLM; this is guaranteed to occur when the initial beliefs are sufficiently close to the MSV REE ones, and thus the intial beliefs about inflation are close to zero.[6]

In order to answer the second question, we will consider agents' initial beliefs that are as far away from the MSV as possible, while still resembling the MSV functional form. To do so, in the next subsection we first derive the Restricted Perceptions Equilibrium, where the 'correct' variable from the MSV is not present at all while the agents forecast using the variable that is irrelevant in the MSV, and then initialize the agents' beliefs in a small neighborhood of the RPE-consistent values.

2.3 Restricted Perception Equilibria

We restrict the agents to use only one endogenous variable in their forecasting models - either a lag of inflation or a lag of output gap. Under the assumption of *i.i.d.* shocks the agents then choose between two models, based on their forecasting performance:

$$\pi_t = \alpha_\pi^\pi + \beta_\pi^\pi \pi_{t-1}, \tag{11}$$

$$\pi_t = \alpha_\pi^y + \beta_\pi^y y_{t-1}, \tag{12}$$

with the coefficients determined by the respective regressions. We call the mis-specified rule in (11) M_π, and the MSV-consistent one in (12) is denoted M_y.[7] The equilibria induced by these forecasting rules are also called M_π and M_y.

The choice of forecasting model influences the dynamics of the model through the expectational term. When the agents are using the 'incorrect' set of variables for forecasting, due to the self-referential

[6] The agents could have included n lags of output gap and k lags of inflation into their PLM, which would have resulted in different strong E-stability conditions.

[7] We call M_y an MSV-consistent rule because it uses the same endogenous variable as the MSV REE. The equilibrium M_y has the same dependence on the lagged output gap as the MSV REE; therefore in what follows we use 'MSV-consistent' and 'MSV REE' interchangeably.

nature of the beliefs the resulting Restricted Perceptions Equilibrium is skewed, and it could happen that the forecasting errors are smaller than they would have been had the agents used the 'correct' variables in this skewed RPE. As the focus of this paper is on initially mis-specified beliefs, we are interested in the conditions under which M_π can be an equilibrium, that is, when this equilibrium is E-stable and the forecasting rule (11) has better ex-post forecasting performance than rule (12).

The Actual Law of Motion (ALM) when the agents are using the M_π forecasting rule is given by the following equations, with the coefficients defined in Appendix D in equations (D73)-(D80):

$$\pi_t = \bar{a}_\pi + \bar{b}_\pi \pi_{t-1} + \bar{c}_\pi y_{t-1} + \bar{\eta}_\pi^g g_t + \bar{\eta}_\pi^u u_t, \tag{13}$$

$$y_t = \bar{a}_y + \bar{b}_y \pi_{t-1} + \bar{c}_y y_{t-1} + \bar{\eta}_y^g g_t + \bar{\eta}_y^u u_t. \tag{14}$$

We measure the forecasting performance of rules M_π and M_y by the mean squared forecast errors (MSFE). We define the error term for MSFE criterion as $E(\pi_t - \hat{\pi}_t)^2$, where π_t is given by the above ALM under the condition that *all* agents' were using (11) to form the expectations. That is, the agents live in equilibrium (13)-(14) induced by the PLM (11), and then an atomistic agent *ex-post* assesses the forecasting performance of alternative rules given the realized inflation over which they have no influence, being measure zero agents.

Proposition 2.0.1 *For the model described in (4-5), M_π equilibrium in (11) is 1) E-stable under condition (D100); 2) M_π produces MSFE that is smaller than M_y MSFE when condition (D105) holds.*

The existence and stability of this RPE is proven in Appendix D.

Until this point the agents did not experience information friction related to the cost of paying attention. We now turn to the attention constraints as in Gabaix (2014) and study how they affect the adaptive learning dynamics and stability.

3 Sparse Rationality under Adaptive Learning

Next, following Audzei and Slobodyan (2022), we initiate the agents' beliefs to be in the small neighborhood of the M_π-consistent ones and allow them to reconsider the variable choices subject to attention constraints à la Gabaix (2014). We are interested in the dynamics of beliefs and the equilibrium they converge to. Will the agents using RLS eventually learn that the coefficient on π is zero, and converge to the MSV-consistent equilibrium? If the answer is Yes, then we are talking about global strong E-stability. As was shown above, when agents do not face attention costs the asymptotic strong E-stability for the MSV REE is obtained when the condition (C60) is satisfied. The numerical analysis in the next section supports the global convergence for zero attention costs.[8] But what happens if the

[8]Outside of the MSV REE, the approximating Strong E-Stability ODE is nonlinear; therefore we cannot prove analytically that the convergence is global.

attention is costly? The agents start very far from the MSV REE beliefs and follow the RLS learning algorithm augmented with the rule for adjusting attention: Will they eventually arrive at the MSV solution? In other words, we want to study the global strong E-stability with positive attention costs.

The Sparse Rationality concept, introduced by Gabaix (2014, 2017) considers the following decision problem. Suppose an agent wants to minimize a loss function associated with some action, with the action being costly. If the loss function is given as a sum of squares of forecast errors from a linear forecasting problem, and the cost of action is a function of the regression coefficients of this regression, this problem is that of *penalized regression*. Depending on the functional form of the penalty, the regression is known as LASSO (the penalty is proportional to the sum of absolute values of coefficients), ridge (the penalty as sum of squared coefficients), or elastic net (a mixture of the two). The penalty term could be interpreted as the sum of costs of paying attention to the variables due to information or data collection efforts.

Gabaix (2014) formulates the decision problem as follows. Suppose that the agents obtain the regression coefficients in their PLM by the usual OLS regression or a linear projection of the response variable y on a set of regressors x:

$$y_t = \widehat{b}_1 x_{1,t-1} + ... + \widehat{b}_n x_{n,t-1} + \epsilon_t. \tag{15}$$

They then form forecasts as $\hat{y_t} = \widehat{b}_1 m_1 x_{1,t-1} + ... + \widehat{b}_n m_n x_{n,t-1}$, with m_i being the *attention weight* allocated to the variable x_i. The agents then maximize the quality of their forecast: $u = -\frac{1}{2} E\left[(\hat{y}_t - y_t)^2\right]$, subject to attention costs $\kappa \sum\limits_{i,j=1...n} |m_i|$. The optimal attention vector m is thus obtained as a solution of the following problem

$$m^* = \arg \min_{m \in [0,1]^n} \left\{ \frac{1}{2} E\left[(\hat{y}_t - y_t)^2\right] + \kappa \sum_{i=1...n} |m_i| \right\}, \tag{16}$$

The problem (16) is known as (non-negative) garrote, and the weights could take any value between 0 and 1. Specifying the penalty term as the sum of absolute values assures that the corner solutions with attention weights of 0 and 1 are possible; this is equivalent to a variable being excluded from and included into the forecasting rule, respectively. Minimizing the loss function with an attention cost penalty is then akin to running a classic model selection exercise; however, it is also possible to pay partial attention to a variable when $0 < m_i < 1$.

We now allow the agents to continue learning adaptively, taking into account the attention cost. As is usual in the adaptive learning literature, they run Recursive Least Squares (RLS) in order to adjust the values of beliefs β and of the second moments of the explanatory variables R, according to equations (17-18):[9]

$$\beta_t = \beta_{t-1} + t^{-1} \cdot R_t^{-1} \cdot z_{t-1} \cdot (T(\beta_{t-1})^T \cdot z_{t-1} + \eta_t - \beta_{t-1}^T \cdot m_{t-1} \cdot z_{t-1})^T, \tag{17}$$
$$R_t = R_{t-1} + t^{-1} \cdot (z_{t-1} z_{t-1}^T - R_{t-1}), \tag{18}$$

with $z_t = [\pi_{t-1}, y_{t-1}]'$, and $\pi_t = T(\beta_{t-1})$ the realized value of inflation (13) given agents beliefs (11).

However, in addition to the standard RLS, they also recompute the optimal attention weights m^*, and adjust their current weights m towards the optimal values, as in equations(19)-(20). We call this combination of usual Adaptive Learning and Sparse Rationality a Sparse AL.

$$m_t = m_{t-1} + nt^{-1} \cdot (m_t^* - m_{t-1}), \tag{19}$$
$$m_t^* = G(\phi_{t-1}, m_{t-1}, R_{t-1}). \tag{20}$$

The last equation describes the solution obtained in (16). The approximating ODE from this procedure, then, is given by the equations (21) below.

$$\begin{aligned} \dot{\beta} &= T(\beta \cdot m) - \beta \cdot m, \\ \dot{R} &= \Sigma - R, \\ \dot{m} &= n \cdot (G(\phi, m, R) - m). \end{aligned} \tag{21}$$

Following Evans and Honkapohja (2001), section 6.2.2, one could interpret this ODE as the agents fixing the parameters of their forecasting rules - β, R, and m - and observing the errors on the right-hand side of (19) for a very large number of periods. Then, they average these right-hand side terms, and make an infinitesimal step in the direction of the average value. Fixing $n = 1$ in (19) then implies that the agents adjust their beliefs, β and R, as often as the attention weights m, while $n < 1$ means a less frequent adjustment of the attention weights than of the beliefs.

4 Dynamics of Sparse AL

In Audzei and Slobodyan (2022) we showed that there exists a large region of parameter values in which the Perceived Law of Motion used by the agents to forecast future values of the inflation and output gap, which includes only the 'wrong' variable's lag, π, induces the M_π RPE. This is so because in this equilibrium the 'wrong' PLM in terms of the forecasting performance could dominate the MSV-consistent PLM with the 'correct' variable y.

[9] Please note that in order to use a shrinkage estimator, we standardize the variables: to demean and divide by their standard deviations. In our dynamic analysis we standardize the variables by dividing by their endogenous mean and standard deviation before calculating weights.

The agents populating this RPE could observe that there was another variable present in the model, namely the output gap y. It further became feasible for them to include this variable into their PLM. However, there were still constraints on the way variables were included into the PLM, expressed as the attention cost in (16).[10] We showed in Audzei and Slobodyan (2022) that solving this penalized regression problem might confirm the agents' initial mis-specified guess, and so they would continue including only the inflation π into their forecasting rule. In terms of the attention vector, the solution was $(m_\pi, m_y) = (x, 0)$, with $1 \geq x > 0$. Therefore, the study concluded that it was possible for the 'wrong' RPE to survive the challenge of a single run of the penalized regression (16) and thus to become self-confirming. However, for some other combinations of penalty and aggressiveness of the Taylor rule, a positive weight on the output gap was indeed optimal.

In this paper we allow the agents to update their beliefs regarding the proper set of variables to be included into the forecasting rule as they run a Sparse AL algorithm, and investigate the convergence properties of the resulting dynamics.

4.1 Convergence to the MSV REE

Figure 1: Attention Weights

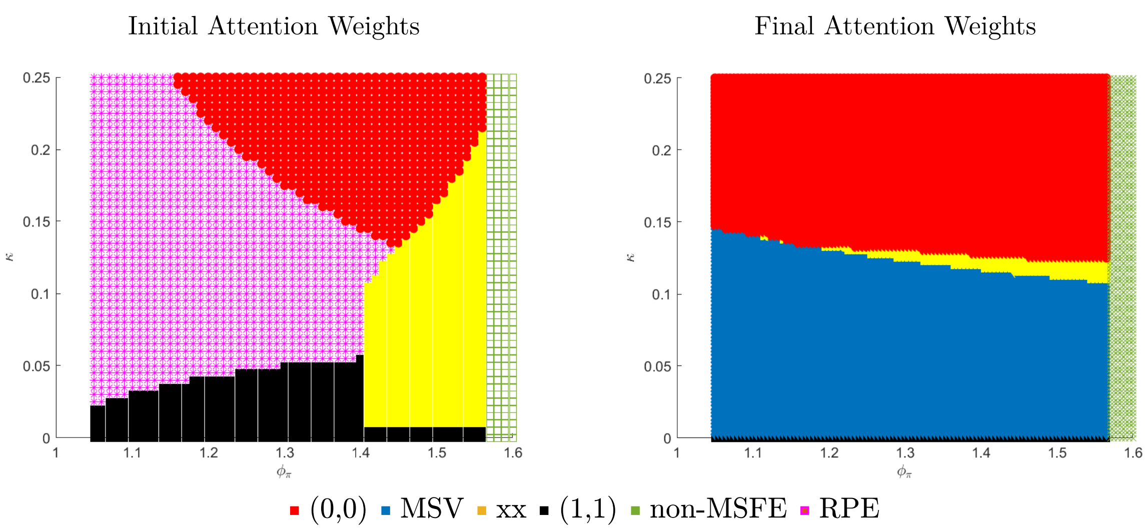


Figure 1 presents the start and end points of such learning dynamics for different values of the attention cost κ and the policy rule's aggressiveness ϕ_π. The left panel presents initial weight, obtained by a single application of the penalized regression (16) in the RPE M_π. When the attention costs are zero, or close to zero, the agents initially choose $(m_\pi, m_y) = (1, 1)$ to pay full attention to both

[10] Audzei and Slobodyan (2022) motivated initial restrictions as limits to process information; they motivated lifting these limitations with an improvement in computing or data collection technologies

variables - black squares in the figures. There is a wide region of the parameter space at the lower values of ϕ_π where the initial weights are consistent with the RPE, $(m_\pi, m_y) = (1, 0)$, pink asterisks. Here, the restricted choice of the variables in the forecasting rule - only π - is reconfirmed by sparse rationality. This region is located at intermediate values of the attention costs and low to medium aggressiveness. For low values of attention cost as well as for intermediate κ combined with higher ϕ_π, both the inflation and output gap tend to have non-zero (yellow area) weights. Also, for even higher attention costs, especially when the Taylor rule is very aggressive, the agents tend to disregard both π and y and forecast inflation as a constant – either long-term inflation or inflation target, both equal to 0 in our model, red area. Finally, the highest values of the aggressiveness by the central bank leads to the RPE M_π forecasting rule producing worse forecasts than the alternative M_y one; thus RPE does not exist in the area of green squares. Notably, the MSV-consistent equilibrium is never an optimal solution. At best, the agents give some weight to the 'correct' variable (output gap), but they never replace the 'wrong' (inflation) with the 'correct' variable. However, this is only a static outcome.

We next let our agents take the coefficients from RPE ALM (13)-(14) as their starting beliefs β and R, and the computed optimal weights as the starting attention weights. Thus, their initial PLM is equivalent to the ALM obtained under the RPE M_π.[11] We then trace the approximating ODE for a sufficiently long time[12] and observe the final beliefs and weights. The final weights are significantly more uniform than the initial ones. The weights converge to the MSV REE (blue area) for a wide range of parameters, to the constant only forecasting rule (red area), or are on the way towards the constant only rule (yellow area). The RPE M_π does not survive for any combination of the parameters (ϕ_π, κ). Naturally, convergence to the constant only rule is observed for high values of κ, while the MSV REE is the limit point for lower values of the attention costs. Convergence to the MSV is observed even for many initial points, with the output gap weight equal to zero. Thus, the agents who are allowed to continue learning, while taking into account the attention penalty, are still able to learn the true equilibrium, unless the attention cost is too high. With stricter monetary policy, agents switch to the constant only rule for smaller κ. The reason for this is lower volatility of inflation and output and lower benefits of paying attention to these variables given attention costs.

Our results do not depend on assuming that the recursive penalized regression learning process starts by taking the RPE M_π ALM as the agents' PLM. Even if the agents take the PLM they hold at the RPE as the starting point, the eventual outcome of learning is still the MSV (0,1) for those pairs (ϕ_π, κ) where we observe convergence to MSV starting from the ALM RPE beliefs. These two exercises suggest that not only could the MSV-consistent equilibrium be strongly E-stable under the attention cost constraint, as we start from over-parametrized PLM, but also that it could be *globally* E-stable in a large region of the parameter space, with the initial beliefs very far from the MSV ones.

[11]f one or both weights are equal to zero, for technical reasons we initiate them with a small positive number ϵ, as otherwise computation of the matrix Σ in (D94) becomes impossible.

[12]We typically use T=30, which in the case of a small constant gain $g = 0.01$ is equivalent to 3,000 periods. In the case of RLS with gains $g_n = 1/n$, continuous time of 30 is equivalent to 1e+13 periods.

In particular, in the second exercise the agents start by having almost zero beliefs in the correct MSV variable y and positive beliefs in the incorrect variable π that is absent in the MSV solution, and still converge to it.

Finally, we have to comment on the relationship between E-stability of the RPE, established by Audzei and Slobodyan (2022), and the strong E-stability of the MSV described above. In the RPE, the agents include *only* the inflation variable into their forecasting rule. The approximating ODE that allowed us to establish E-stability of the RPE is thus 1-dimensional in the beliefs (β) space. In contrast, once the agents start learning subject to attention costs, they explicitly take into account that there could be two variables in their PLM, and thus the approximating ODE is 2-dimensional in β space. In addition, there are dynamics in the m space which weren't present in the analysis of RPE E-stability. In other words, the nature of the dynamic adjustment of beliefs (and attention weights) changes dramatically, in particular through expansion of dimensionality. Even if the initial belief on output in the agents' PLM is zero, they still *could* move in that dimension, while during convergence to the RPE the β_y dimension didn't exist.

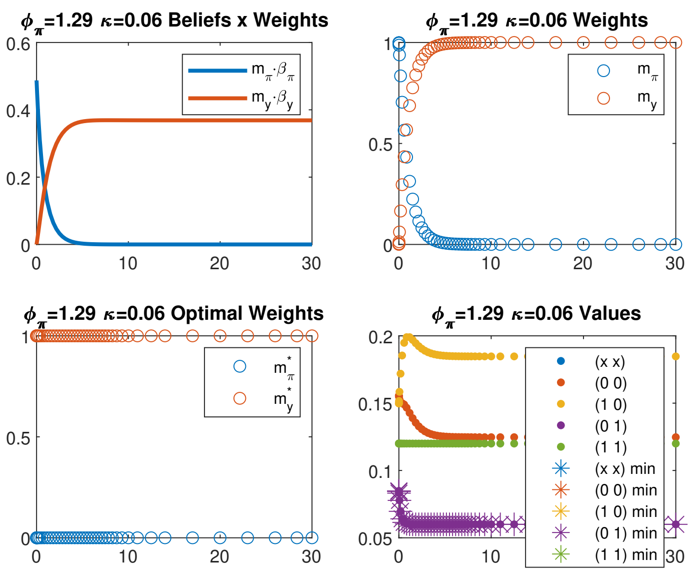
We now turn to discussion of the thin yellow wedge on the final weights figure, which exhibits rather non-trivial dynamics with switching of the forecasting rule's functional form.

4.2 Sliding dynamics

The approximating ODE trajectories converging to the red (constant only) or blue area (MSV) are rather simple in terms of the optimal attention weights evaluated along the trajectory: they are either (0,0) or (0,1) throughout the whole trajectory, respectively. In this case, the actual attention weights are adjusted monotonically to their limit values. In particular along the trajectory converging to the MSV REE, the weight on inflation m_π declines while the weight on output gap m_y is monotonically increasing. The lower left panel of Figure 2 shows the optimal weights, and the upper right panel the actual attention weights that the agents hold while they are adjusting towards the $(0, 1)$ limit. Asymptotically the total impact of the output gap on the agents' inflation forecast, $m_y \cdot \beta_y$, upper left panel, is the same as the corresponding value at the MSV REE despite the attention costs. The agents pay the costs but still prefer to use the correct forecasting rule and the correct coefficient in it. The lower right panel shows that the corner solution $(0, 1)$ remains optimal (violet asterisks) throughout the whole trajectory.

Similar dynamics are observed for most trajectories that converge to the $(0, 0)$ limit weights: the corner solution $(0, 0)$ remains optimal for the whole duration of the simulation. However, there are other types of trajectories that exhibit a switching behavior of the optimal solution in the attention weights space, and we now turn to the detailed discussion of these solutions. These are the trajectories converging to the yellow circles and yellow diamonds, see Figure 3 for an example. At some point along the trajectory around t=2.5, the value of the objective function in (16) obtained for the MSV-

Figure 2: Convergence of weights and beliefs



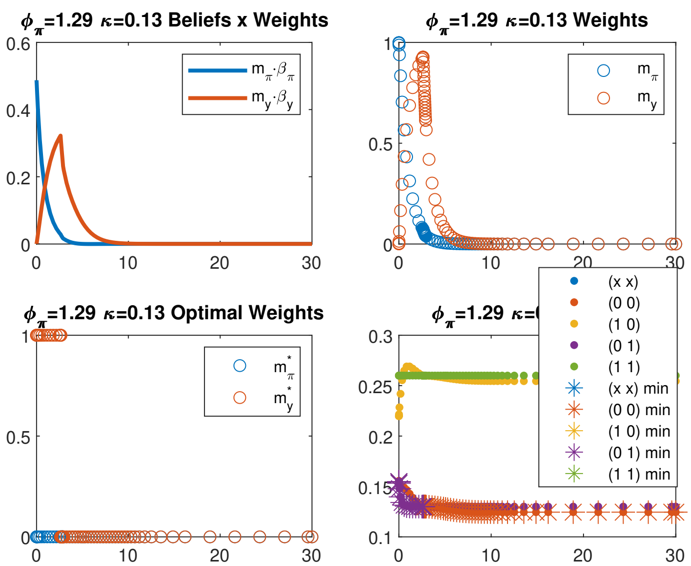
The figure illustrates the convergence of weights to MSV REE consistent values for attention costs $\kappa = 0.06$ and monetary policy reaction to inflation $\phi_\pi = 1.29$. Continuous time units of the approximating ODE are on the horizontal axis. Convergence to MSV REE.

consistent weights, V(0,1), becomes equal to the value generated by the constant only weights, V(0,0). After this point, the monotonic convergence to the MSV weights is replaced with a convergence to the constants only rule $(0, 0)$. The switch is best seen in the lower right panel of the Figure 3: Before $t \approx 2.5$ it is the (0,1) solution that produces the minimal value (violet asterisks), but after this time it's the (0,0) solution which becomes the best (orange asterisks).

The points where such a situation happens form a surface in the (β, R, m) space. At one side of the surface, we have V(0,0)>V(0,1), while at the other side the opposite situation takes place. The optimal weights, respectively, are (0,1) and (0,0). From the ODE for attention weights (21) we then see that the right-hand side in the equation for m_y is discontinuous at this boundary. Importantly, it could happen that due to this discontinuity the flow described by (21) points back to the boundary on *both* sides of it, making it intuitively clear that locally the ODE trajectories will be attracted to the boundary.

In order to study the dynamics in this case, we need to turn to the theory of non-smooth differential equations described in Appendix E. As is described there, a *sliding* dynamic could occur along the boundary on which the two solutions to the problem (16) give exactly the same value. This happens when the flow described by (21) points in the direction of the boundary on both sides of it, making

Figure 3: Convergence of weights and beliefs



The figure illustrates the convergence of weights to $(m_\pi, m_y) = (0, 0)$ for selected values of attention costs $\kappa = 0.13$ and monetary policy reaction to inflation $\phi_\pi = 1.29$. Continuous time units of the approximating ODE are on the horizontal axis. Convergence to constant only forecasting rule.

possible a stable trajectory that lies entirely within the boundary for some time interval.

In our case, we observed several types of the trajectories encountering the boundary between the solutions $(m_\pi, m_y) = (0, 0)$ and $(0, 1)$. Two were the most common. The first type encounters the boundary, punches through it (the scalar product of projections of the flow on the normal to the boundary from two sides is positive), and continues evolving according to the ODE (21) towards the constant only forecasting rule. These are the yellow diamonds in Figure 1. Another type encounters the boundary and settles into the sliding dynamics as the scalar product of projections on the normal is negative, eventually converging to the constant only forecasting rule. All points represented by the yellow circles in Figure 1 denote such dynamics. Occasionally, a trajectory that first punched through the boundary, then encountered it for a second time and settled for the sliding dynamics, was observed. We also encountered a few trajectories whereby the sliding dynamics ended before time T and the trajectory then continued along the non-boundary ODE (21). Up to three episodes of sliding could occur along the convergence trajectory for some parameter values.

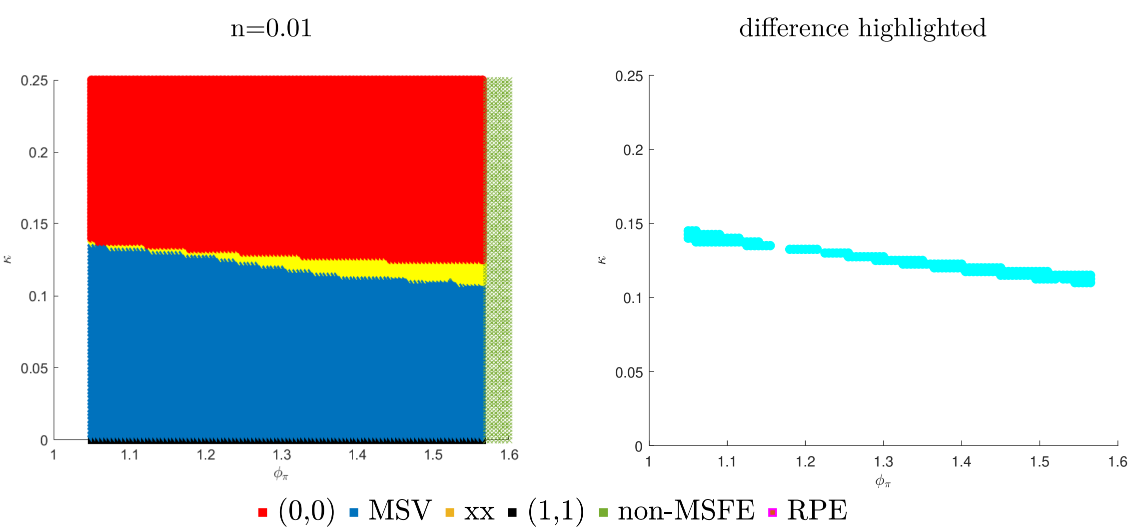
Importantly, no trajectories that encountered the boundary were observed to converge to the MSV REE attention weights, whether or not the trajectory was converging to the MSV before the encounter.

4.3 The Slow Weights Learning Case and Non-Smooth Dynamics

The case $n = 1$ assumes that the agents update attention weights with the same speed as their OLS beliefs. One, however, could entertain different hypotheses. Generally speaking, reconsideration of the set of variables to be included into the forecasting rule and of the attention weights for different variables is a significantly more complex task than updating R, computing its inverse and multiplying it by the forecast error to get the iteration of β. Determining optimal weights is a constrained optimization problem that in a multi-dimensional case requires comparison of multiple corner solutions. Therefore, it may make sense for the agents to reconsider their weights less frequently than their OLS beliefs. This would then amount to $n \ll 1$ in the updating equations (20-21).

We present the results of the relatively slow learning of attention weights ($n = 0.01$) in Figure 4. The left panel shows the final weights, while on the right we show the difference with the results for $n = 1$. Blue points are the parameters' values for which in the case $n = 1$ we had convergence to the MSV REE, but with $n = 0.01$ the constants only forecasting rule is the final outcome. All of these trajectories that have switched the final attention weights are those that encounter the boundary between (0,1) and (0,0) solutions along the way. In order to understand this behavior we look carefully into a simplified version of the model dynamics.

Figure 4: Attention Weights



In order to generate the intuition for the results on the slow sliding dynamics we inspect the equations (21) and see that what matters for the dynamics of both beliefs and attention weights are the element-wise products of attention weight and beliefs $\beta * m$, which are the total effects of the two variables ($m_x \cdot \beta_x$) on the overall inflation forecast. For the trajectories where the optimal attention weight on inflation, m_π, remains equal to 0, only the impact $\Psi_y = m_y \cdot \beta_y$ matters. Therefore, in order to simplify the exposition we switch our attention to the dynamics in a two-dimensional space $\underline{x} = (m_y, \beta_y)$.[13] The ODE (21) in this space is given by the following equations:

$$\begin{aligned} \dot{\beta}_y &= \bar{c}_\pi - \beta_y \cdot m_y, \\ \dot{m}_y &= n \cdot \left(m_y^* - m_y\right). \end{aligned} \tag{22}$$

In this space, the boundary between the two corner solutions (0,0) and (0,1) is given as the solution to the equation $V(0,0) = V(0,1)$ which is $\Psi_y = m_y \cdot \beta_y = \bar{\Psi}_y$, a hyperbola in the two-dimensional space (m_y, β_y). Using notation from Appendix E, the equation for the boundary is

$$\sigma(m_y, \beta_y) = m_y \cdot \beta_y - \bar{\Psi}_y = 0.$$

The discontinuity above and below the boundary comes from the fact that the optimal solution for m_y is either 0 or 1 at different sides of it. For $\sigma(m_y, \beta_y)$ below the boundary, $m_y^* = 1$ while above the boundary $m_y^* = 0$.[14] Write the time derivative of σ as

$$\begin{aligned} \dot{\sigma} = (m_y \dot{\cdot} \beta_y) &= \dot{m}_y \cdot \beta_y + \dot{\beta}_y \cdot m_y \\ &= \left(\bar{c}_\pi - \bar{\Psi}_y\right) \cdot m_y + n \cdot \left(m_y^* - m_y\right). \end{aligned}$$

The first term in the last line is always positive, while the second is negative above the boundary, where $m_y^* = 0$, and positive below it, because $m_y^* = 1$. When the second term is larger in absolute value than the first, the boundary is stable, as $\dot{\sigma}$ is negative for $\sigma > 0$ and positive for $\sigma < 0$. Sliding dynamics ensue. However, when we decrease n, the second term becomes smaller in the absolute value. It is now possible to have $\dot{\sigma} > 0$ also for $\sigma > 0$, and there is no sliding as the boundary is simply punched through.

With sliding, the system evolves along the boundary $\sigma(\underline{x}) = 0$. Given that the time derivative of β_y is a positive constant at the boundary, the value of β_y grows without bounds during sliding. However, as the product of β_y and m_y at the boundary is constant, m_y must converge to zero. Therefore, the limit point of the sliding dynamics in this simple case could only be $(m_\pi, m_y) = (0, 0)$. This behavior is probably responsible for the fact that once the sliding dynamics commences in our simulations that take place in 7D space, the constant only solution (0,0) is the ultimate outcome, even when the sliding is consequently discontinued: sliding brings the trajectory ever closer to (0,0) rather than back to the (0,1) solution, the MSV REE.

Another consequence of the slow updating of attention weights consists of affecting whether the

[13] We further assume that the second moments R have converged to their equilibrium values Σ in order to simplify the exposition.

[14] This is because the value of the penalty term is increasing in m_y, and so the forecasting rule with fewer variables is preferred when we increase m_y marginally from the boundary.

trajectory even reaches the (0,0)-(0,1) boundary. Outside of the simplified 2D case we just considered, the boundary is a complicated object in the seven-dimensional space rather than a simple hyperbola $\Psi_y = m_y \cdot \beta_y = \bar{\Psi}_y$. It is possible that when the trajectory is moving towards $m_y = 1$ very fast ($n = 1$) hitting the boundary becomes impossible, thus expanding the region in the parameter space where convergence to the MSV REE is observed.

5 Conclusions

In this paper we extend the standard Recursive Least Square learning algorithm to the case of penalized regression as in Gabaix (2014). We investigate the convergence properties of the continuous time approximating ODE for this combined algorithm, called Sparse Adaptive Learning, and establish that allowing for dynamic choice of attention to be paid to different model variables rules out convergence to the RPE. The attention weights corresponding to the RPE are never the ultimate outcome, even though initially the beliefs are consistent with the RPE. This result is in stark contrast with a single application of the sparsity penalized regression, which never delivered the MSV REE as an outcome in Audzei and Slobodyan (2022). Depending on the attention costs, a limit point of the learning dynamics is either MSV REE or the constant-only rule.

Strictness of the monetary policy affects the evolution of the learning algorithm. When the Taylor rule is more aggressive, the model variables become less volatile and less correlated across time making lags of endogenous variables less useful for forecasting. In the presence of attention costs this could lead to the agents selecting a constants-only forecasting rule – long-term inflation or inflation target – rather than the rule consistent with the MSV REE.

The learning algorithm considered in this paper could lead to non-smooth dynamics due to the agents discontinuously selecting the set of variables to be included into their forecasting rule. The presence of these non-smooth dynamics forces us to rely on the theory of non-smooth differential equations to study the approximating ODE. We also establish that the relative speed of adjusting the belief coefficients and the attention weights has important implications for the trajectories that could encounter the boundary between the two corner solutions, and develop some analytical results in a simplified case.

The asymptotic global E-Stability of MSV REE we demonstrate in the paper implies that even when the agents who are allowed to reconsider choices of their forecasting rule in a self-referential system subject to attention costs initiate learning from the 'wrong' equilibrium, they would typically learn the MSV REE; alternatively, in a system with little volatility or autocorrelation of the endogenous variables, they will switch to using the constant-only rule.

References

Adam, K., 2005. Learning to forecast and cyclical behavior of output and inflation. Macroeconomic Dynamics 9, 1–27.

Adam, K., 2007. Experimental evidence on the persistence of output and inflation. Economic Journal 117, 603–636.

Andrle, M., Bruha, J., 2023. A Sparse Kalman Filter: A Non-Recursive Approach. Working Papers 2023/13. Czech National Bank. URL: https://ideas.repec.org/p/cnb/wpaper/2023-13.html.

Ascari, G., Mavroeidis, S., McClung, N., 2023. Coherence without rationality at the zero lower bound. Journal of Economic Theory 214, 105745. URL: https://www.sciencedirect.com/science/article/pii/S0022053123001412, doi:https://doi.org/10.1016/j.jet.2023.105745.

Audzei, V., 2023. Learning and cross-country correlations in a multi-country DSGE model. Economic Modelling 120. URL: https://ideas.repec.org/a/eee/ecmode/v120y2023ics0264999322003790.html, doi:10.1016/j.econmod.2022.10.

Audzei, V., Slobodyan, S., 2022. Sparse restricted perceptions equilibrium. Journal of Economic Dynamics and Control 139. doi:10.1016/j.jedc.2022.10441.

Banbura, M., Giannone, D., Modugno, M., Reichlin, L., 2013. Now-casting and the real-time data flow, in: Elliott, G., Granger, C., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier. volume 2 of *Handbook of Economic Forecasting*. chapter 0, pp. 195–237.

Branch, W.A., Evans, G.W., 2006. Intrinsic heterogeneity in expectation formation. Journal of Economic Theory 127, 264–295.

Branch, W.A., McGough, B., Zhu, M., 2022. Statistical sunspots. Theoretical Economics 17, 291–329.

Bullard, J., Mitra, K., 2002. Learning about monetary policy rules. Journal of Monetary Economics 49, 1105 – 1129.

De Mol, C., Giannone, D., Reichlin, L., 2008. Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? Journal of Econometrics 146, 318–328.

Evans, G.W., Honkapohja, S., 2001. Learning and expectations in macroeconomics. Princeton University Press, Princeton.

Evans, G.W., Honkapohja, S., Sargent, T., Williams, N., 2012. Bayesian model averaging, learning and model selection. CDMA Working Paper Series 201203. Centre for Dynamic Macroeconomic Analysis.

Filippov, A.F., 1988. Differential Equations with Discontinuous Righthand Sides. 1 ed., Springer. doi:https://doi.org/10.1007/978-94-015-7793-9.

Gabaix, X., 2014. A sparsity-based model of bounded rationality. The Quarterly Journal of Economics 129, 1661–1710.

Gabaix, X., 2017. Behavioral inattention. Working Paper 24096. National Bureau of Economic Research.

Galí, J., 2015. Monetary policy, inflation, and the business cycle: an introduction to the New Keynesian framework and its applications. 2 ed., Princeton University Press.

Gefang, D., 2014. Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage. International Journal of Forecasting 30, 1–11.

Gibbs, C.G., 2017. Forecast combination, non-linear dynamics, and the macroeconomy. Economic Theory 63, 653–686. doi:10.1007/s00199-016-0951-x.

Hajdini, I., 2022. Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model. Working Papers 22-03R. Federal Reserve Bank of Cleveland. URL: `https://ideas.repec.org/p/fip/fedcwq/93720.html`, doi:10.26509/frbc-wp-202203r.

Hansen, C., Liao, Y., 2019. The factor-lasso and k-step bootstrap approach for inference in high-dimensional economic applications. Econometric Theory 35, 465–509.

Hommes, C., Zhu, M., 2014. Behavioral learning equilibria. Journal of Economic Theory 150, 778–814.

Hommes, C.H., 2014. Behaviorally rational expectations and almost self-fulfilling equilibria. Review of Behavioral Economics 1, 75–97.

Jeffrey, M., 2019. Hidden Dynamics: The Mathematics of Switches, Decisions and Other Discontinuous Behaviour. 1 ed., Springer. doi:10.1007/978-3-030-02107-8.

Korobilis, D., 2013. Hierarchical shrinkage priors for dynamic regressions with many predictors. International Journal of Forecasting 29, 43–59.

McCallum, B.T., 1983. On non-uniqueness in rational expectations models : An attempt at perspective. Journal of Monetary Economics 11, 139–168.

McCallum, B.T., 2003. The unique minimum state variable RE solution is E-stable in all well formulated linear models. NBER Working Papers 9960. National Bureau of Economic Research, Inc.

Mele, A., Molnár, K., Santoro, S., 2020. On the perils of stabilizing prices when agents are learning. Journal of Monetary Economics 115, 339–353. URL: https://ideas.repec.org/a/eee/moneco/v115y2020icp339-353.html, doi:10.1016/j.jmoneco.2019.08.

Molnar, K., 2007. Learning with expert advice. Journal of the European Economic Association 5, 420–432.

Nazemi, A., Fabozzi, F.J., 2018. Macroeconomic variable selection for creditor recovery rates. Journal of Banking and Finance 89, 14–25.

Ormeno, A., Molnar, K., 2015. Using survey data of inflation expectations in the estimation of learning and rational expectations models. Journal of Money, Credit and Banking 47, 673–699.

Orphanides, A., Williams, J.C., 2007. Robust monetary policy with imperfect knowledge. Journal of Monetary Economics 54, 1406–1435.

Pfajfar, D., Žakelj, B., 2014. Experimental evidence on inflation expectation formation. Journal of Economic Dynamics and Control 44, 147–168.

Piiroinen, P.T., Kuznetsov, Y.A., 2008. An event-driven method to simulate filippov systems with accurate computing of sliding motions. ACM Trans. Math. Softw. 34. URL: https://doi.org/10.1145/1356052.1356054, doi:10.1145/1356052.1356054.

Slobodyan, S., Bogomolova, A., Kolyuzhnov, D., 2016. Stochastic gradient learning and instability: an example. Macroeconomic Dynamics 20, 777–790.

Slobodyan, S., Wouters, R., 2012a. Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models. American Economic Journal: Macroeconomics 4, 65–101. URL: https://ideas.repec.org/a/aea/aejmac/v4y2012i2p65-101.html.

Slobodyan, S., Wouters, R., 2012b. Learning in an estimated medium-scale DSGE model. Journal of Economic Dynamics and Control 36, 26–46.

Stock, J., Watson, M., 2016. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics, in: Taylor, J.B., Uhlig, H. (Eds.), Handbook of Macroeconomics. Elsevier. volume 2 of *Handbook of Macroeconomics*. chapter 0, pp. 415–525.

Stock, J.H., Watson, M.W., 2002. Macroeconomic Forecasting Using Diffusion Indexes. Journal of Business and Economic Statistics 20, 147–162.

Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Vázquez, J., Aguilar, P., 2021. Adaptive learning with term structure information. European Economic Review 134. URL: https://ideas.repec.org/a/eee/eecrev/v134y2021ics0014292121000428.html, doi:10.1016/j.euroecorev.2021.

Yuan, M., Lin, Y., 2007. On the non-negative garrotte estimator. Journal of the Royal Statistical Society Series B 69, 143–161.

A Model Derivations

Households maximize the infinite discounted sum of utility over consumption C_t and labour decisions N_t:

$$\sum_{t=0}^{\infty} \beta^t U(C_t, H_t, N_t), \tag{A23}$$

where $H_t = hC_{t-1}$ is external habit and $0 < \beta < 1$ is the discount factor. The optimization results in the familiar conditions:

$$-\frac{U_{n,t}}{U_{c,t}} = \frac{W_t}{P_t}, \tag{A24}$$

$$Q_t = E_t \left[\frac{U_{c,t+1}}{U_{c,t}} \frac{P_t}{P_{t+1}} \right]. \tag{A25}$$

The first equation equalizes agent's utility of consumption and dis-utility of labour. The second is an Euler equation determining agents' inter-temporal consumption decisions. We assume utility separable in consumption and labour, with σ - relative utility of risk aversion, ϕ - Frisch elasticity of labour supply, and g_t a preference shock:

$$U_t = e^{g_t} \left(\frac{(C_t - H_t)^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\phi}}{1+\phi} \right) \tag{A26}$$

so that

$$U_{c,t} = e^{g_t}(C_t - H_t)^{-\sigma}. \tag{A27}$$

Plugging $Y_t = C_t$ into the linearized Euler equation, we get the investment-savings curve (1).

For the rest of model, we utilize a textbook model from Galí (2015). The firms use labour to produce differentiated final goods and face nominal rigidities á la Calvo with the probability of optimizing a price θ. The differentiated good is aggregated using a consumption index: $C_t = \left[\int_0^1 C_t(i)^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}$. Firms' pricing decisions result in a new Keynesian Phillips curve in (2) with $\omega \equiv (1-\theta)(1-\beta\theta)(\sigma+\phi)/\theta$.

It is convenient to rewrite the system of equation (4):(5) as:

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = AE_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} + C \begin{bmatrix} \pi_{t-1} \\ y_{t-1} \end{bmatrix} + B \begin{bmatrix} u_t \\ g_t \end{bmatrix}, \tag{A28}$$

with $A = \begin{bmatrix} \beta - \frac{\omega(1-h)(\phi_\pi - 1)}{(1+h)\sigma} & \frac{\omega}{1+h} \\ -\frac{1-h}{(1+h)\sigma}(\phi_\pi - 1) & \frac{1}{1+h} \end{bmatrix}$, $C = \begin{bmatrix} 0 & \frac{\omega h}{(1+h)} \\ 0 & \frac{h}{(1+h)} \end{bmatrix}$, $B = \begin{bmatrix} 1 & \frac{\omega(1-h)}{(1+h)\sigma} \\ 0 & \frac{1-h}{(1+h)\sigma} \end{bmatrix}$.

B Expert Forecast of the Output Gap

We first derive output gap expectations, given the inflation PLM, which can be potentially inconsistent with REE MSV solution. Suppose inflation PLM has the following form

$$\pi_t = \psi_\pi \pi_{t-1} + \psi_y y_{t-1}, \tag{B29}$$

$$E_t \pi_{t+1} = \psi_\pi \pi_t + \psi_y y_t = \psi_\pi^2 \pi_{t-1} + \psi_\pi \psi_y y_{t-1} + \psi_y y_t, \tag{B30}$$

Plugging this PLM into (4) - (5) and denoting $b_\pi^\pi = \left(\beta - \frac{\omega(1-h)(\phi_\pi-1)}{(1+h)\sigma}\right)$ and $b_\pi^y = -\frac{1-h}{(1+h)\sigma}(\phi_\pi - 1)$, we get:

$$y_t = \frac{1}{1-b_\pi^y\psi_y}\left(b_\pi^y\psi_\pi^2\pi_{t-1} + (b_\pi^y\psi_\pi\psi_y + \frac{h}{1+h})y_{t-1} + \frac{1}{1+h}E_t y_{t+1} + \frac{1-h}{(1+h)\sigma}g_t\right), \tag{B31}$$

Similarly, for inflation:

$$\begin{aligned}\pi_t = &\frac{b_\pi^\pi\psi_\pi^2}{1-b_\pi^y\psi_y}\pi_{t-1} + \frac{\omega+\beta\psi_y}{(1-b_\pi^y\psi_y)(1+h)}E_t y_{t+1}\\ &+\left(\frac{b_\pi^\pi\psi_\pi\psi_y}{1-b_\pi^y\psi_y} + \frac{h(\omega+\psi_y\beta)}{(1+h)(1-b_\pi^y\psi_y)}\right)y_{t-1} + \frac{(1-h)(\omega+\psi_y\beta)}{(1+h)\sigma(1-b_\pi^y\psi_y)}g_t + u_t,\end{aligned} \tag{B32}$$

where we have used $b_\pi^\pi - \omega b_\pi^y = \beta$.

We further assume that the agents receive ”expert advice” which coincides with the MSV solution for output given the current ALM.[15]. Thus the expert forecasting model is:

$$y_t = \gamma_y y_{t-1} + \gamma_\pi\pi_{t-1} + \gamma_g g_t + \gamma_u u_t, \tag{B33}$$

$$y_{t+1} = \gamma_y y_t + \gamma_\pi\pi_t = \gamma_y^2 y_{t-1} + \gamma_\pi\gamma_y\pi_{t-1} + \gamma_\pi\pi_t + \gamma_y\gamma_g g_t + \gamma_y\gamma_u u_t. \tag{B34}$$

Now, as these experts know the system in (B31)-(B32), they plug the expression for inflation and rearrange using $b_\pi^\pi - \omega b_\pi^y = \beta$:

$$\begin{aligned}&E_t y_{t+1}((1-b_\pi^y\psi_y)(1+h) - \gamma_\pi(\omega+\beta\psi_y))\\ &= (\gamma_\pi\psi_y\psi_\pi b_\pi^\pi(1+h) + \gamma_\pi h(\omega+\psi_y\beta) + (h+1)\gamma_y^2\,(1-b_\pi^y\psi_y))y_{t-1}\\ &+\gamma_\pi(1+h)(\gamma_y(1-b_\pi^y\psi_y) + b_\pi^\pi\psi_\pi^2)\pi_{t-1}+\\ &(\gamma_y\gamma_g(1-b_\pi^y\psi_y)(1+h) + \frac{\gamma_\pi(1-h)}{\sigma}(\beta\psi_y+\omega))g_t\\ &+(1+h)(1-b_\pi^y\psi_y)(\gamma_y\gamma_u+\gamma_\pi)u_t.\end{aligned} \tag{B35}$$

Now, we redefine the coefficients such that

$$E_t y_{t+1} = \tilde{\gamma}_y y_{t-1} + \tilde{\gamma}_\pi\pi_{t-1} + \tilde{\gamma}_u u_t + \tilde{\gamma}_g g_t, \tag{B36}$$

with

$$\tilde{\gamma}_y = \frac{(\gamma_\pi\psi_y\psi_\pi b_\pi^\pi(1+h) + \gamma_\pi h(\omega+\psi_y\beta) + (h+1)\gamma_y^2\,(1-b_\pi^y\psi_y))}{((1-b_\pi^y\psi_y)(1+h) - \gamma_\pi(\omega+\beta\psi_y))}, \tag{B37}$$

$$\tilde{\gamma}_\pi = \frac{\gamma_\pi(1+h)(\gamma_y(1-b_\pi^y\psi_y) + b_\pi^\pi\psi_\pi^2)}{((1-b_\pi^y\psi_y)(1+h) - \gamma_\pi(\omega+\beta\psi_y))}, \tag{B38}$$

$$\tilde{\gamma}_u = \frac{(1+h)(1-b_\pi^y\psi_y)(\gamma_y\gamma_u+\gamma_\pi)}{((1-b_\pi^y\psi_y)(1+h) - \gamma_\pi(\omega+\beta\psi_y))}, \tag{B39}$$

$$\tilde{\gamma}_g = \frac{(\gamma_y\gamma_g(1-b_\pi^y\psi_y)(1+h) + \frac{\gamma_\pi(1-h)}{\sigma}(\beta\psi_y+\omega))}{((1-b_\pi^y\psi_y)(1+h) - \gamma_\pi(\omega+\beta\psi_y))}, \tag{B40}$$

[15]That is, taking into account agents' PLM for inflation.

which will be our expert advice.

To get the coefficients, we plug the expression into (B31). The coefficients of the experts' rule will be the solution to the following equations and are functions of agents' PLM.

$$\gamma_\pi = \frac{b^y_\pi \psi^2_\pi}{(1-b^y_\pi\psi_y)} + \frac{\gamma_\pi(\gamma_y(1-b^y_\pi\psi_y)+b^\pi_\pi\psi^2_\pi)}{(1-b^y_\pi\psi_y)((1-b^y_\pi\psi_y)(1+h)-\gamma_\pi(\omega+\beta\psi_y))}, \tag{B41}$$

$$\gamma_y = \frac{b^y_\pi\psi_\pi\psi_y(1+h)+h}{(1-b^y_\pi\psi_y)(1+h)} + \frac{1}{(1-b^y_\pi\psi_y)(1+h)}\frac{(\gamma_\pi\psi_y\psi_\pi b^\pi_\pi(1+h)+\gamma_\pi h(\omega+\psi_y\beta)+(h+1)\gamma_y^2\left(1-b^y_\pi\psi_y\right))}{((1-b^y_\pi\psi_y)(1+h)-\gamma_\pi(\omega+\beta\psi_y))}, \tag{B42}$$

$$\gamma_g = \frac{h-1}{\sigma\left((h+1)\left(b^y_\pi\psi_y-1\right)+\gamma_\pi\left(\beta\psi_y+\omega\right)+\gamma_y\right)}, \tag{B43}$$

$$\gamma_u = \frac{\gamma_\pi}{(h+1)\left(1-b^y_\pi\psi_y\right)-\gamma_\pi\left(\beta\psi_y+\omega\right)-\gamma_y}. \tag{B44}$$

C Rational Expectations MSV

Under REE MSV, the perceived law of motion for the system in (A28) is:

$$\begin{bmatrix}\pi_t\\ y_t\end{bmatrix} = \Omega + \bar{C}\begin{bmatrix}\pi_{t-1}\\ y_{t-1}\end{bmatrix} + \Gamma\begin{bmatrix}u_t\\ g_t\end{bmatrix}, \tag{C45}$$

$$E_t\begin{bmatrix}\pi_{t+1}\\ y_{t+1}\end{bmatrix} = \Omega + \bar{C}\begin{bmatrix}\pi_t\\ y_t\end{bmatrix} = \Omega + \bar{C}\Omega + \bar{C}^2\begin{bmatrix}\pi_{t-1}\\ y_{t-1}\end{bmatrix} + \bar{C}\Gamma\begin{bmatrix}u_t\\ g_t\end{bmatrix}. \tag{C46}$$

Plugging the PLM into (A28):

$$\begin{bmatrix}\pi_t\\ y_t\end{bmatrix} = A(I+\bar{C})\Omega + \left[A\left(\bar{C}\right)^2 + C\right]\begin{bmatrix}\pi_{t-1}\\ y_{t-1}\end{bmatrix} + \left[A\bar{C}\Gamma + B\right]\begin{bmatrix}u_t\\ g_t\end{bmatrix}. \tag{C47}$$

Using the method of undetermined coefficients we can solve for the PLM coefficients from:

$$\bar{C} = A\bar{C}^2 + C = 0, \tag{C48}$$

$$\Gamma = B + A\bar{C}\Gamma, \tag{C49}$$

$$\Omega = A(I+\bar{C})\Omega; \tag{C50}$$

Generically, the eigenvalues of $A(I+\bar{C})$ are not equal to unity, and therefore the solution for the constant vector Ω is a zero vector.

The MSV coefficients are the solution of the following system

$$c^y_\pi = \frac{\omega h}{1+h} + \left(\beta - \frac{\omega(1-h)(\phi_\pi-1)}{(1+h)\sigma}\right)c^y_\pi c^y_y + \frac{\omega}{1+h}{c^y_y}^2, \tag{C51}$$

$$c^y_y = \frac{h}{1+h} - \frac{(1-h)(\phi_\pi-1)}{(1+h)\sigma}c^y_\pi c^y_y + \frac{1}{1+h}{c^y_y}^2. \tag{C52}$$

While the first equation is quadratic, the second is cubic, with the determinant changing signs depending on the central bank's reaction function. Thus we have 3 possible solutions. Following McCallum (1983), we choose the solution for c^y_π which goes to zero when $\omega = 0$.

Multiplying the second equation by ω and subtracting from the first, we get an expression for c_π^y:

$$c_\pi^y = \frac{\omega c_y^y}{1-\beta c_y^y}. \tag{C53}$$

For $0 < c_y^y < 1$ it follows form (D83), that $c_\pi^y > 0$.

We define the solution as $\bar{C} = \begin{bmatrix} 0 & c_\pi^y \\ 0 & c_y^y \end{bmatrix}$, $\Gamma = \begin{bmatrix} 1 & \gamma_\pi^y \\ 0 & \gamma_y^y \end{bmatrix}$, $\Omega = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

C.1 E-Stability

We understand E-stability as expectational stability under learning as described in Evans and Honkapohja (2001). For our model, the concept could be illustrated for our model as follows. Given the solution, one can write T-mapping as a system of equations:

$$c_\pi^y \quad -> \quad \frac{\omega h}{1+h} + \left(\beta - \frac{\omega(1-h)(\phi_\pi - 1)}{(1+h)\sigma}\right) c_\pi^y c_y^y + \frac{\omega}{1+h} {c_y^y}^2, \tag{C54}$$

$$c_y^y \quad -> \quad \frac{h}{1+h} - \frac{(1-h)(\phi_\pi - 1)}{(1+h)\sigma} c_\pi^y c_y^y + \frac{1}{1+h} {c_y^y}^2, \tag{C55}$$

with the Jacobian

$$\begin{bmatrix} (\beta - \frac{\omega(1-h)(\phi_\pi-1)}{(1+h)\sigma}) c_y^y - 1 & (\beta - \frac{\omega(1-h)(\phi_\pi-1)}{(1+h)\sigma}) c_\pi^y + \frac{2\omega}{1+h} c_y^y \\ -\frac{(1-h)(\phi_\pi-1)}{(1+h)\sigma} c_y^y & -\frac{(1-h)(\phi_\pi-1)}{(1+h)\sigma} c_\pi^y - 1 + \frac{2}{1+h} c_y^y \end{bmatrix} = \begin{bmatrix} J11 & J12 \\ J21 & J22 \end{bmatrix}. \tag{C56}$$

As long as the Taylor principle $\phi_\pi > 1$ is satisfied, $J11$ and $J21$ are negative. For reasonable ϕ_π, $J12 > 0$.[16]

For $J22 < 0$, the following conditions must hold: $c_y^y \leq (1+h)/2$ and $\phi_\pi > 1$. For $(1+h)/2 \geq c_y^y \leq 1$, the condition for even stricter $\phi_\pi > 1 + \frac{2c_y^y - (1+h)}{(1-h)c_\pi^y} > 1$.

The discriminant (product of eigenvalues) of (C56) is then $-J12J21 + J11J22$ and the trace (sum of eigenvalues) is $J11 + J22$. The discriminant is positive and the trace is negative. Thus, the sufficient condition for both eigenvalues to be negative is the Taylor principle $\phi_\pi > 1$. The Taylor principle is sufficient unless the process for output is highly persistent $c_y^y \geq (1+h)/2$. For a highly persistent output process, the sufficient condition for E-stability is $\phi_\pi > 1 + \frac{2c_y^y - (1+h)}{(1-h)c_\pi^y} > 1$.

The **strong E-Stability** is stability under learning when agents are allowed to have over-specified forecasting rules. In our model, we allow for arbitrary matrix $\bar{C}$ with the following coefficients:

$$\bar{C} = \begin{bmatrix} v & c_\pi^y \\ w & c_y^y \end{bmatrix}. \tag{C57}$$

The part of T-mapping responsible for $\bar{C}$ is modified:

$$\begin{bmatrix} b_\pi^\pi & \frac{\omega}{1+h} \\ b_\pi^y & \frac{1}{1+h} \end{bmatrix} \begin{bmatrix} v & c_\pi^y \\ w & c_y^y \end{bmatrix} \begin{bmatrix} v & c_\pi^y \\ w & c_\pi^y \end{bmatrix} + \begin{bmatrix} 0 & \frac{\omega h}{1+h} \\ 0 & \frac{h}{1+h} \end{bmatrix} =$$

$$\begin{bmatrix} b_\pi^\pi (v^2 + w c_\pi^y) + \frac{\omega w}{1+h}(v + c_y^y) & b_\pi^\pi (v + c_y^y) c_\pi^y + \frac{\omega}{1+h}(w c_\pi^y + (c_y^y)^2) + \frac{\omega h}{1+h} \\ b_\pi^y (v^2 + w c_\pi^y) + \frac{w}{1+h}(v + c_y^y) & b_\pi^y (v + c_y^y) c_\pi^y + \frac{h}{1+h}(w c_\pi^y + (c_y^y)^2) + \frac{h}{1+h} \end{bmatrix}, \tag{C58}$$

[16] To make $J12$ negative, the reaction of monetary policy to inflation should be stronger than empirically plausible.

where the elements of matrices A and C are given in (A28).

The part of the Jacobian responsible for these coefficients becomes:

$$J = \begin{bmatrix} 2b_\pi^\pi \bar{v} + \frac{\omega}{1+h}\bar{w} - 1 & b_\pi^\pi \bar{c}_\pi^y + \frac{\omega}{1+h}(\bar{v} + \bar{c}_y^y) & b_\pi^\pi \bar{w} & \frac{\omega}{1+h}\bar{w} \\ 2b_\pi^y \bar{v} + \frac{1}{1+h}\bar{w} & b_\pi^y \bar{c}_\pi^y + \frac{1}{1+h}(\bar{v} + \bar{c}_y^y) - 1 & b_\pi^y \bar{w} & \frac{1}{1+h}\bar{w} \\ b_\pi^\pi \bar{c}_\pi^y & \frac{\omega}{1+h}\bar{c}_\pi^y & b_\pi^\pi(\bar{v} + \bar{c}_y^y) + \frac{\omega}{1+h}\bar{w} - 1 & b_\pi^\pi \bar{c}_\pi^y + 2\frac{\omega}{1+h}\bar{c}_y^y \\ b_\pi^y \bar{c}_\pi^y & \frac{1}{1+h}\bar{c}_\pi^y & b_\pi^y \bar{c}_y^y & b_\pi^y \bar{c}_\pi^y + 2\frac{1}{1+h}\bar{c}_y^y - 1 \end{bmatrix} \tag{C59}$$

The condition for strong E-Stability is

$$Eig(J) < 0. \tag{C60}$$

With $\bar{v} = \bar{w} = 0$, the Jacobian becomes:

$$J = \begin{bmatrix} -1 & b_\pi^\pi \bar{c}_\pi^y + \frac{\omega}{1+h}\bar{c}_y^y & 0 & 0 \\ 0 & b_\pi^y \bar{c}_\pi^y + \frac{1}{1+h}\bar{c}_y^y - 1 & 0 & 0 \\ b_\pi^\pi \bar{c}_\pi^y & \frac{\omega}{1+h}\bar{c}_\pi^y & b_\pi^\pi \bar{c}_y^y - 1 & b_\pi^\pi \bar{c}_\pi^y + 2\frac{\omega}{1+h}\bar{c}_y^y \\ b_\pi^y \bar{c}_\pi^y & \frac{1}{1+h}\bar{c}_\pi^y & b_\pi^y \bar{c}_y^y & b_\pi^y \bar{c}_\pi^y + 2\frac{1}{1+h}\bar{c}_y^y - 1 \end{bmatrix}, \tag{C61}$$

where the lower 2×2 block is the same as (C56) for weak E-stability. In addition to two eigenvalues identical to those of (C56), this matrix has two more eigenvalues: -1 and $b_\pi^y \bar{c}_\pi^y + \frac{1}{1+h}\bar{c}_y^y - 1$. For the second extra eigenvalue to be negative, $b_\pi^y \bar{c}_\pi^y + \frac{1}{1+h}\bar{c}_y^y - 1 = J22 - 1 < J22$ must hold.

Thus, the sufficient condition for strong E-stability is satisfied as long as the sufficient condition for weak E-stability is satisfied.

D Restricted Perception Equilibrium

D.1 Definition of RPE and ALM coefficients

We focus on the RPE M_π and derive the conditions for its existence below.

The agents' inflation expectations with the M_π forecasting rule (11) are formulated as follows:

$$\pi_{t+1} \quad = \quad \alpha_\pi^\pi + \beta_\pi^\pi \pi_t = \alpha_\pi^\pi(1 + \beta_\pi^\pi) + (\beta_\pi^\pi)^2 \pi_{t-1}, \tag{D62}$$

When we plug the above M_π into the model in (5) using expert forecast for output (B36), we get the actual laws of motion (ALM) for inflation:

$$\begin{aligned} \pi_t &= b_\pi^\pi \alpha_\pi^\pi (1 + \beta_\pi^\pi) + \left(b_\pi^\pi (\beta_\pi^\pi)^2 + \frac{\tilde{\gamma}_\pi \omega}{1+h} \right) \pi_{t-1} + \frac{\omega(h + \tilde{\gamma}_y)}{1+h} y_{t-1} \\ &+ \frac{\omega((1-h) + \sigma\tilde{\gamma}_g)}{(1+h)\sigma} g_t + (\frac{\omega\tilde{\gamma}_u}{1+h} + 1)u_t = \\ &= \bar{a}_\pi + \bar{b}_\pi \pi_{t-1} + \bar{c}_\pi y_{t-1} + \bar{\eta}_\pi^g g_t + \bar{\eta}_\pi^u u_t, \end{aligned} \tag{D63}$$

with $b_\pi = \beta - \frac{\omega(\phi_\pi - 1)(1-h)}{\sigma(1+h)}$, and $\bar{x}_\pi$ being the coefficients in inflation ALM.

Similarly, the ALM for the output gap is:

$$\begin{aligned} y_t = b_\pi^y (\alpha_\pi^\pi (1 + \beta_\pi^\pi)) + \left(b_\pi^y (\beta_\pi^\pi)^2 + \frac{\tilde{\gamma}_\pi}{1+h} \right) \pi_{t-1} + \frac{h + \tilde{\gamma}_y}{1+h} y_{t-1} + \frac{1 - h + \sigma\tilde{\gamma}_g}{(1+h)\sigma} g_t + \\ \frac{\tilde{\gamma}_u}{1+h} u_t = \bar{a}_y + \bar{b}_y \pi_{t-1} + \bar{c}_y y_{t-1} + \bar{\eta}_y^g g_t + \bar{\eta}_y^u u_t, \end{aligned} \tag{D64}$$

with $b^y_\pi = -\frac{1-h}{(1+h)\sigma}(\phi_\pi - 1)$, and $\bar{x}_y$ are output gap ALM coefficients.

When focusing on a converged RPE M_π, we can plug $\psi_y = 0$ and $\psi_\pi = \beta^\pi_\pi$ into expert forecast coefficients. The solution for γ and $\tilde{\gamma}$ is the solution to the following equations:

$$\gamma_\pi = b^y_\pi(\beta^\pi_\pi)^2 + \frac{\gamma_\pi(\gamma_y + b^\pi_\pi(\beta^\pi_\pi)^2)}{1+h-\gamma_\pi\omega}, \tag{D65}$$

$$\gamma_y = \frac{h+\gamma_y^2}{1+h-\gamma_\pi\omega}, \tag{D66}$$

$$\gamma_g = \frac{1-h}{\sigma\left(1+h-\gamma_\pi\omega-\gamma_y\right)}, \tag{D67}$$

$$\gamma_u = \frac{\gamma_\pi}{((1+h)-\gamma_\pi\omega)-\gamma_y}; \tag{D68}$$

and

$$\tilde{\gamma}_y = \frac{\gamma_\pi h\omega + (1+h)\gamma_y^2}{1+h-\gamma_\pi\omega}, \tag{D69}$$

$$\tilde{\gamma}_\pi = \frac{\gamma_\pi(1+h)(\gamma_y + b^\pi_\pi(\beta^\pi_\pi)^2)}{1+h-\gamma_\pi\omega}, \tag{D70}$$

$$\tilde{\gamma}_u = \frac{(1+h)\gamma_\pi}{((1+h)-\gamma_\pi\omega-\gamma_y)}, \tag{D71}$$

$$\tilde{\gamma}_g = (1-h)\frac{\gamma_y+\omega\gamma_\pi}{\left(\sigma\left(1+h-\gamma_\pi\omega-\gamma_y\right)\right)}. \tag{D72}$$

It is instructive to examine the coefficients. An economically meaningful coefficient on lagged output is $0 < \gamma_y < 1$. It follows from $(D66)$ that $\omega\gamma_\pi \le 1+h-2\sqrt{h}$ and $\gamma_\pi < (1-\gamma_y^2)/\omega$; and $0 < \tilde{\gamma}_y < 1$.

Plugging (D65):(D72) into (D63):(D64), we obtain ALM coefficients for output and inflation.

$$\bar{c}_\pi = \omega\frac{(h+\tilde{\gamma}_y)}{1+h} = \omega\gamma_y, \tag{D73}$$

$$\bar{b}_\pi = b^\pi_\pi(\beta^\pi_\pi)^2 + \frac{\tilde{\gamma}_\pi\omega}{1+h} = \omega\gamma_\pi + \beta(\beta^\pi_\pi)^2, \tag{D74}$$

$$\bar{c}_y = \frac{h+\tilde{\gamma}_y}{1+h} = \gamma_y, \tag{D75}$$

$$\bar{b}_y = b^y_\pi(\beta^\pi_\pi)^2 + \frac{\tilde{\gamma}_\pi}{1+h} = \gamma_\pi. \tag{D76}$$

Now, the ALM coefficients for shock processes:

$$\bar{\eta}^g_\pi = \frac{(1-h)\omega}{(1+h-\gamma_\pi\omega-\gamma_y)\sigma}, \tag{D77}$$

$$\bar{\eta}^u_\pi = \frac{(1+h)-\gamma_y}{((1+h)-\gamma_\pi\omega-\gamma_y)}, \tag{D78}$$

$$\bar{\eta}_y^g = \frac{1-h}{(1+h-\gamma_\pi\omega-\gamma_y)\sigma}, \tag{D79}$$

$$\bar{\eta}_y^u = \frac{\gamma_\pi}{1+h-\gamma_\pi\omega-\gamma_y}. \tag{D80}$$

D.2 RPE Beliefs

We treat the agents as econometricians, who learn the coefficients from running regressions of the corresponding PLMs. Denoting the covariance between inflation and output $Cov(\pi, y) \equiv \sigma_{\pi y}$, and the variances of output and inflation as σ_y^2 and σ_π^2 respectively, we can derive the coefficients for the M_π forecasting rule:

$$\begin{aligned}
\beta_\pi^\pi &= \frac{Cov(\pi_t, \pi_{t-1})}{Var(\pi_{t-1})} = \frac{Cov(\bar{b}_\pi \pi_{t-1} + \bar{c}_\pi y_{t-1}, \pi_{t-1})}{Var(\pi_{t-1})} = \\
&= \bar{b}_\pi + \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}, \qquad \text{(D81)} \\
\alpha_\pi^\pi &= (1-\beta_\pi^\pi)\bar{\pi}. \qquad \text{(D82)}
\end{aligned}$$

For the M_y rule, the regression coefficients are computed with the law of motion for inflation given by M_π in (D63):

$$\begin{aligned}
\beta_y^y &= \frac{Cov(\pi_t, y_{t-1})}{Var(y_{t-1})} = \frac{Cov\left(\bar{b}_\pi \pi_{t-1} + \bar{c}_\pi y_{t-1}, y_{t-1}\right)}{Var(y_{t-1})} = \\
&= \bar{b}_\pi \frac{\sigma_{\pi y}}{\sigma_y^2} + \bar{c}_\pi, \qquad \text{(D83)} \\
\alpha_y^y &= (1-\beta_y^y)\bar{\pi}. \qquad \text{(D84)}
\end{aligned}$$

D.3 Existence of RPE

For the M_π to exist, there must exist a β_π^π, which is a solution for (D81). It has been shown in Audzei and Slobodyan (2022) that there exists a unique solution for (D81), as long as the following matrix D is stable:

$$D = \begin{bmatrix} \bar{b}_\pi & \bar{c}_\pi \\ \bar{b}_y & \bar{c}_y \end{bmatrix} = \begin{bmatrix} \omega\gamma_\pi + \beta(\beta_\pi^\pi)^2 & \omega\gamma_y \\ \gamma_\pi & \gamma_y \end{bmatrix}. \tag{D85}$$

Determinant and trace are given by the following expressions:

$$\begin{aligned}
det(D) &= -\gamma_\pi\omega\gamma_y + (\omega\gamma_\pi + \beta(\beta_\pi^\pi)^2)\gamma_y = (\beta_\pi^\pi)^2\gamma_y\beta \qquad \text{(D86)} \\
tr(D) &= \omega\gamma_\pi + \beta(\beta_\pi^\pi)^2 + \gamma_y \qquad \text{(D87)}
\end{aligned}$$

For the matrix to be stable, we use the following conditions (see Audzei and Slobodyan 2022, Appendix B, for details):

$$det(D) < 1, \tag{D88}$$

$$det(D) > tr(D) - 1, \tag{D89}$$

$$det(D) > -tr(D) - 1. \tag{D90}$$

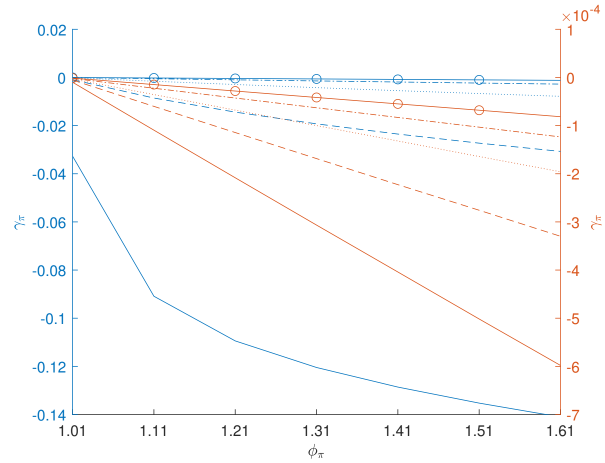


Figure 5: γ_π for different values of r – relative variance of inflationary shocks. The figure is drawn for r in the range [0.1:0.1:0.5] left axis and [0.6:0.1:1] right axis.

The condition in (D88) is satisfied as $(\beta_\pi^\pi)^2\gamma_y\beta < 1$.

To prove that the (D89) is satisfied, we combine (D86) and (D87) and rewrite them as:

$$\omega\gamma_\pi + \beta(\beta_\pi^\pi)^2 + \gamma_y < (\beta_\pi^\pi)^2\gamma_y\beta + 1 \tag{D91}$$

For $\gamma_\pi \leq 0$, given that $\beta_\pi^\pi < 1$, $\beta < 1$, and $\gamma_y < 1$, it is straightforward to show that $\beta(\beta_\pi^\pi)^2 + \gamma_y < 1 + \beta(\beta_\pi^\pi)^2\gamma_y$.

Values of $\gamma_\pi > 0$ are not economically meaningful: besides, during our simulations there was no stable solution with $\gamma_\pi > 0$ for our parametrization. In Figure 5, we plot the solutions for γ_π as a function of monetary policy response to inflation and relative volatility of mark-up shocks to show that the solution for γ_π is always below zero.

The condition (D90) is satisfied as long as $(\beta_\pi^\pi)^2\gamma_y\beta > 0$ and $\gamma_\pi < 0$.

Thus, the matrix D is stable and the unique RPE solution exists.

D.4 M_π Mapping and Variance-Covariance Matrix

To calculate observed average inflation, rewrite M_π ALM as

$$\left[I - \begin{pmatrix} \bar{b}_\pi & \bar{c}_\pi \\ \bar{b}_y & \bar{c}_y \end{pmatrix}\right]\begin{bmatrix}\bar{\pi} \\ \bar{y}\end{bmatrix} = \begin{bmatrix}\bar{a}_\pi \\ \bar{a}_y\end{bmatrix}, \tag{D92}$$

$$\begin{bmatrix}\bar{\pi} \\ \bar{y}\end{bmatrix} = \begin{bmatrix}\frac{\bar{a}_y\bar{c}_\pi + \bar{a}_\pi(1-\bar{c}_y)}{1-(\bar{b}_\pi+\bar{c}_y)+(\bar{b}_\pi\bar{c}_y-\bar{b}_y\bar{c}_\pi)} \\ \frac{\bar{a}_y-\bar{a}_y\bar{b}_\pi+\bar{a}_\pi\bar{b}_y}{1-(\bar{b}_\pi+\bar{c}_y)+(\bar{b}_\pi\bar{c}_y-\bar{b}_y\bar{c}_\pi)}\end{bmatrix} \tag{D93}$$

To calculate the variances-covariance matrix of the ALM, re-write:

$$\Sigma = \begin{bmatrix} \sigma_\pi^2 & \sigma_{\pi y} \\ \sigma_{\pi y} & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} \bar{b}_\pi & \bar{c}_\pi \\ \bar{b}_y & \bar{c}_y \end{bmatrix} \begin{bmatrix} \sigma_\pi^2 & \sigma_{\pi y} \\ \sigma_{\pi y} & \sigma_y^2 \end{bmatrix} \begin{bmatrix} \bar{b}_\pi & \bar{b}_y \\ \bar{c}_\pi & \bar{c}_y \end{bmatrix} +$$
$$+ \begin{bmatrix} \bar{\eta}_\pi^u & \bar{\eta}_\pi^g \\ \bar{\eta}_y^u & \bar{\eta}_y^g \end{bmatrix} \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_g^2 \end{bmatrix} \begin{bmatrix} \bar{\eta}_\pi^u & \bar{\eta}_y^u \\ \bar{\eta}_\pi^g & \bar{\eta}_y^g \end{bmatrix} =,$$
$$= \begin{bmatrix} (\bar{b}_\pi)^2\sigma_\pi^2 + 2\bar{b}_\pi\bar{c}_\pi\sigma_{\pi y} + (\bar{c}_\pi)^2\sigma_y^2 & \bar{b}_\pi\bar{b}_y\sigma_\pi^2 + (\bar{c}_\pi\bar{b}_y + \bar{b}_\pi\bar{c}_y)\sigma_{\pi y} + \bar{c}_\pi\bar{c}_y\sigma_y^2 \\ \bar{b}_\pi\bar{b}_y\sigma_\pi^2 + (\bar{c}_\pi\bar{b}_y + \bar{b}_\pi\bar{c}_y)\sigma_{\pi y} + \bar{c}_\pi\bar{c}_y\sigma_y^2 & (\bar{c}_y)^2\sigma_y^2 + (\bar{b}_y)^2\sigma_\pi^2 + 2\bar{b}_y\bar{c}_y\sigma_{\pi y} \end{bmatrix} +$$
$$+ \begin{bmatrix} (\bar{\eta}_\pi^u)^2\sigma_u^2 + (\bar{\eta}_\pi^g)^2\sigma_g^2 & \bar{\eta}_\pi^g\bar{\eta}_y^g\sigma_g^2 + \bar{\eta}_\pi^u\bar{\eta}_y^u\sigma_u^2 \\ \bar{\eta}_\pi^g\bar{\eta}_y^g\sigma_g^2 + \bar{\eta}_\pi^u\bar{\eta}_y^u\sigma_u^2 & (\bar{\eta}_y^g)^2\sigma_g^2 + (\bar{\eta}_y^u)^2\sigma_u^2 \end{bmatrix}. \tag{D94}$$

The elements of the variance-covariance matrix are the solution of the following equations:

$$\sigma_\pi^2 = (\bar{b}_\pi)^2\sigma_\pi^2 + 2\bar{b}_\pi\bar{c}_\pi\sigma_{\pi y} + (\bar{c}_\pi)^2\sigma_y^2 + (\bar{\eta}_\pi^u)^2\sigma_u^2 + (\bar{\eta}_\pi^g)^2\sigma_g^2, \tag{D95}$$
$$\sigma_{\pi y} = \bar{b}_\pi\bar{b}_y\sigma_\pi^2 + (\bar{c}_\pi\bar{b}_y + \bar{b}_\pi\bar{c}_y)\sigma_{\pi y} + \bar{c}_\pi\bar{c}_y\sigma_y^2 + \bar{\eta}_\pi^g\bar{\eta}_y^g\sigma_g^2 + \bar{\eta}_\pi^u\bar{\eta}_y^u\sigma_u^2, \tag{D96}$$
$$\sigma_y^2 = (\bar{c}_y)^2\sigma_y^2 + (\bar{b}_y)^2\sigma_\pi^2 + 2\bar{b}_y\bar{c}_y\sigma_{\pi y} + (\bar{\eta}_y^g)^2\sigma_g^2 + (\bar{\eta}_y^u)^2\sigma_u^2. \tag{D97}$$

D.5 E-stability

Proof of Proposition 2.0.1

From (D81) and (D82), the T-map for the RPE M_π is:

$$\beta_\pi^\pi \to \bar{b}_\pi + \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}, \tag{D98}$$
$$\alpha_\pi^\pi \to (1-\beta_\pi^\pi)\bar{\pi}. \tag{D99}$$

E-Stability For M_π to by E-stable, eigenvalues of the following matrix should be negative:

$$Eig \begin{bmatrix} (1-\beta_\pi^\pi)\frac{\partial(\bar{\pi})}{\partial\alpha_\pi^\pi} - 1 & \frac{\partial((1-\beta_\pi^\pi)\bar{\pi})}{\partial\beta_\pi^\pi} \\ 0 & \frac{\partial\left[\bar{b}_\pi + \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}\right]}{\partial\beta_\pi^\pi} - 1 \end{bmatrix} =$$
$$\begin{bmatrix} (1-\beta_\pi^\pi)\frac{\partial(\bar{\pi})}{\partial\alpha_\pi^\pi} - 1 & \frac{\partial\left[\bar{b}_\pi + \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}\right]}{\partial\beta_\pi^\pi} - 1 \end{bmatrix} < 0. \tag{D100}$$

In the text we have assumed that $\bar{\pi} = 0$, in this case the first eigenvalue is negative. We plot the second eigenvalue for the parameter range $r = \sigma_u/\sigma_g \in (0:1]$ for $\phi_\pi > 1$. As both eigenvalues are negative for the considered parameter range, we conclude that M_π is E-Stable.

D.6 Better forecasting performance

For better forecasting performance of M_π relative to M_y we consider the mean squared forecast errors criterion, given that agents have previously selected M_π and test alternative models given M_π ALM in (D63). It is convenient to denote the composite of shocks as $\mu_t \equiv \bar{\eta}_\pi^g g_t + \bar{\eta}_\pi^u u_t$. We start with the

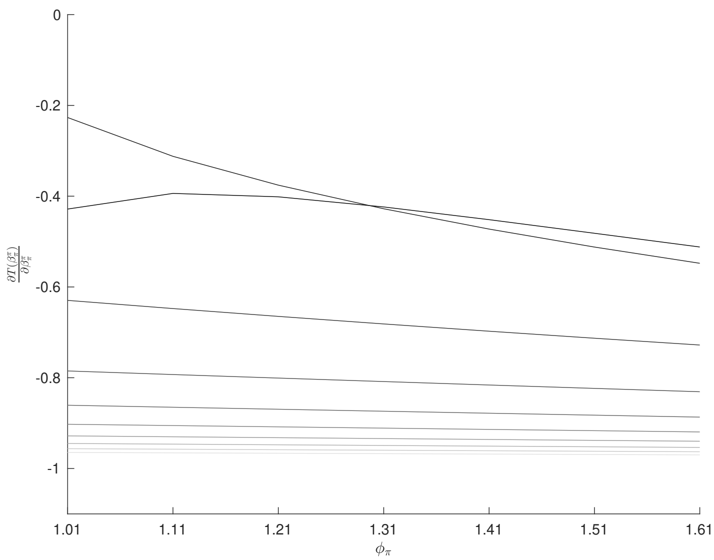


Figure 6: E-Stability. Note: The figure is drawn for r in the range [0.1 : 0.1 : 1]. The darker colours correspond to smaller r - lower relative standard deviations of inflationary shocks.

mean forecast error of M_π. The forecast error of M_π is the difference between the forecast and actual inflation:

$$
\begin{aligned}
: \quad e_t^\pi &= (\beta_\pi^\pi - \bar{b}_\pi)\pi_{t-1} - \bar{c}_\pi y_{t-1} + \mu_t \\
&= (\bar{b}_\pi + \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2} - \bar{b}_\pi)\pi_{t-1} - \bar{c}_\pi y_{t-1} - \mu_t \\
&= \bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}\pi_{t-1} - \bar{c}_\pi y_{t-1} + \mu_t, \qquad \text{(D101)}
\end{aligned}
$$

$$
\begin{aligned}
&: \quad MSFE_\pi = E_t(e_t^\pi)^2 = E_t[\bar{c}_\pi \frac{\sigma_{\pi y}}{\sigma_\pi^2}(\pi_{t-1}) - \bar{c}_\pi(y_{t-1}) + \mu_t]^2 \\
&: \quad = E_t[\bar{c}_\pi^2 (\frac{\sigma_{\pi y}}{\sigma_\pi^2})^2(\pi_{t-1})^2 - 2\bar{c}_\pi R \bar{c}_\pi(\pi_{t-1})(y_{t-1}) + \bar{c}_\pi^2(y_{t-1})^2 + \mu^2] \\
&: \quad = \bar{c}_\pi^2 \sigma_y^2 (1 - \frac{\sigma_{\pi y}^2}{\sigma_\pi^2 \sigma_y^2}) + \sigma_\mu^2, \qquad \text{(D102)}
\end{aligned}
$$

Similarly, the forecast error of M_y is:

$$
\begin{aligned}
&: \quad e_t^y = \left(c_y^y - \bar{c}_\pi\right) y_{t-1} - \bar{b}_\pi \pi_{t-1} + \mu_t = \\
&: \quad = \bar{b}_\pi \sigma_{\pi y} y_{t-1} - \bar{b}_\pi \pi_{t-1} + \mu_t = \\
&: \quad MSFE_y = E[\bar{b}_\pi[\frac{\sigma_{\pi y}}{\sigma_y^2}(y_{t-1}) - (\pi_{t-1})]^2 + \mu^2] \qquad \text{(D103)} \\
&: \quad = \bar{b}_\pi^2 \sigma_\pi^2 [1 - \frac{\sigma_{\pi y}^2}{\sigma_y^2 \sigma_\pi^2}] + \sigma_\mu^2. \qquad \text{(D104)}
\end{aligned}
$$

We are looking for the conditions under which $MSFE_\pi < MSFE_y$. Then, the criterion is simply:

$$\bar{c}_\pi^2 \sigma_y^2 < \bar{b}_\pi^2 \sigma_\pi^2. \tag{D105}$$

E Theoretical Foundations of Sliding Dynamics

The discussion in this section follows Jeffrey (2019), Ch. 2, and the concepts from Filippov (1988).

Suppose there's a vector ODE with a discontinuous flow,

$$\dot{\underline{\mathbf{x}}} = f(\underline{\mathbf{x}}, \lambda), \tag{E106}$$

so that at the boundary defined by $\mathcal{D} = \{\underline{\mathbf{x}} : \sigma(\underline{\mathbf{x}}) = 0\}$ there is a discontinuity of the function f. The surface $\mathcal{D}$ is called *discontinuity surface*. The *switching multiplier* λ could be selected so that $\lambda = sign(\sigma)$. Denote

$$\begin{aligned} f^+(\underline{\mathbf{x}}) \quad &: \quad = f(\underline{\mathbf{x}}; +1),\ \sigma(\underline{x}) > 0, \\ f^-(\underline{\mathbf{x}}) \quad &: \quad = f(\underline{\mathbf{x}}; -1),\ \sigma(\underline{x}) < 0. \end{aligned}$$

Then the time derivative of the flow above (below) the surface could be written as

$$\frac{d}{dt} = \frac{d\underline{\mathbf{x}}}{dt}\frac{d}{d\underline{\mathbf{x}}} = f^{\pm}\frac{d}{d\underline{\mathbf{x}}}.$$

The normal vector to $\mathcal{D}$ is defined as $\frac{d\sigma}{d\underline{\mathbf{x}}}$. Then, $f \cdot \frac{d\sigma}{d\underline{\mathbf{x}}} = \frac{d\underline{\mathbf{x}}}{dt}\frac{d\sigma}{d\underline{\mathbf{x}}} = \frac{d\sigma}{dt} = \dot{\sigma}$, so the projection of the vector f onto the normal vector to $\mathcal{D}$ gives the time derivative of σ.

The Lemma 2.1 of Jeffrey (2019) then states that if $f(\underline{\mathbf{x}}, \lambda)$ is continuous in λ and the components of $f^{\pm}(\underline{\mathbf{x}})$ normal to the boundary are opposing to each other, there exists an intermediate value of λ, denoted $\lambda^{\$}$, $-1 \leq \lambda^{\$} \leq 1$, such that $f\left(\underline{\mathbf{x}}; \lambda^{\$}\right) \cdot \frac{d\sigma}{d\underline{\mathbf{x}}} = 0$. One could then further define solutions of the ODE (E106) that exist on the discontinuity surface, the *sliding* flow, so that

$$\begin{aligned} \dot{\underline{\mathbf{x}}} = f^{\$}(\underline{\mathbf{x}}) = f\left(\underline{\mathbf{x}}; \lambda^{\$}\right) \text{ for } \sigma(\underline{x}) = 0, \\ f\left(\underline{\mathbf{x}}; \lambda^{\$}\right) \cdot \frac{d\sigma}{d\underline{\mathbf{x}}} = 0. \end{aligned} \tag{E107}$$

This flow's projection onto the normal to the boundary equals zero; thus $\sigma(\underline{\mathbf{x}}) = 0$ is preserved over time. However, there could be a non-zero projection to the subspace that is tangent to the boundary $\mathcal{D}$ at the point where it is reached by the original flow. This projection tangential to $\mathcal{D}$ gives rise to the sliding dynamics along the boundary.

If the components of $f^{\pm}(\underline{\mathbf{x}})$ normal to the boundary are pointing in the same direction, then a simple *crossing* of the boundary will happen, and no sliding along the boundary $\mathcal{D}$ will be observed. We check the opposing condition by computing the scalar product of the flows f^+ and f^-, with a negative value signifying oppositely directed projections and thus the presence of sliding.

The easiest way of generating a function that is smooth in λ is to postulate that

$$\begin{aligned} &\dot{\underline{\mathbf{x}}} = f(\underline{\mathbf{x}}, \lambda) = \frac{1}{2}(1+\lambda) f^+(\underline{\mathbf{x}}) + \frac{1}{2}(1-\lambda) f^-(\underline{\mathbf{x}}), \\ &\lambda = +1,\ \sigma(\underline{x}) > 0, \\ &\lambda = -1,\ \sigma(\underline{x}) < 0. \end{aligned} \tag{E108}$$

Then, one could define λ^s so that the projection of $\frac{1}{2}(1+\lambda^s) f^+(\underline{\mathbf{x}}) + \frac{1}{2}(1-\lambda^s) f^-(\underline{\mathbf{x}})$ on the normal to the boundary $\mathcal{D}$ is zero. The resulting flow then produces trajectories that slide along the boundary.

The construction above then suggests the following simple algorithm of evaluating the trajectories of the approximating ODE that could involve sliding dynamics.

1. Trace trajectory of the ODE solution until time T, stopping at $min(\tau, T)$, where τ is the first time the boundary V(0,0)=V(0,1) is reached.

2. If $\tau < T$, numerically compute the normal to the boundary V(0,0)=V(0,1) at the point it is achieved.

3. Check whether the scalar product of the projections of the flow $f^{\pm}(\underline{\mathbf{x}})$ onto the normal to the boundary is positive or negative.

4. If the product is positive, this is a simple crossing. Continue with Step 1, stopping at $min(\tau^*, T)$, where τ^* is the next time the boundary V(0,0)=V(0,1) is reached. Otherwise, switch to simulating the sliding ODE constructed as in (E107-E108) above, also until $min(\tau^*, T)$.

5. If $\tau^* < T$, repeat Step 3, otherwise end.

The algorithm described above could be thought of as a simplified version of Piiroinen and Kuznetsov (2008).

Abstrakt

Tento článek studuje konvergenční vlastnosti, včetně lokální a globální silné E-stability, rovnováhy racionálních očekávání při nehladké dynamice učení. V jednoduchém novokeynesiánském modelu uvažujeme dva typy informačních omezení působících společně – adaptivní učení a řídkou racionalitu. Pro různá počáteční přesvědčení (beliefs) zkoumáme, zda konvergence k minimální stavové proměnné rovnováhy racionálních očekávání (MSV REE) dochází v průběhu času při kladných nákladech na pozornost. Zjišťujeme, že pro jakákoli počáteční přesvědčení prognostické pravidlo agentů konverguje buď k rovnováze MSV REE, nebo pro velké náklady na pozornost k pravidlu, které nebere v úvahu všechny proměnné kromě konstanty. Přísnější měnová politika mírně zvýhodňuje pravidlo pouze s konstantou. Chybně specifikované prognostické pravidlo, které používá proměnnou, jež se v MSV REE nevyskytuje, tento algoritmus nepřežije. Teorie nehladkých diferenciálních rovnic je použita ke studiu dynamiky našeho algoritmu učení.

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: https://www.cerge-ei.cz/

Editor: Byeongju Jeong

The paper is available online at https://www.cerge-ei.cz/working-papers/.

ISBN 978-80-7343-599-8 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)