



Discussion Paper Series

Heterogeneous learning dynamics and speed of convergence

By

Michele Berardi

Centre for Growth and Business Cycle Research, Economic Studies,
University of Manchester, Manchester, M13 9PL, UK

August 2010
Number 148

Download paper from:

<http://www.socialsciences.manchester.ac.uk/cgbcr/discussionpapers/index.html>

Heterogeneous learning dynamics and speed of convergence

Michele Berardi
The University of Manchester

August 2010

Abstract

In a simple, forward looking linear stochastic model we investigate the impact of heterogeneity in expectations formation on the speed of convergence of the learning process of agents towards equilibrium. We find that even when heterogeneity does not affect learnability in term of its asymptotic outcome, it can still have an important impact on the learnability of an equilibrium in terms of the speed of convergence of learning dynamics.

Keywords: Heterogeneous expectations, learning, speed of convergence.

JEL Codes: *C62, D83.*

1 Introduction

The literature on learning in macroeconomics proposes to step back from the rational expectations assumption and to represent agents as adaptive learners that update over time their beliefs about the economy using specific expectations formation procedures. A typical way to model this feature is to endow agents with a forecasting model and a least squares recursive learning algorithm, and evaluate whether agents expectations converge over time to an equilibrium point. See for example Marcet and Sargent (1989) and Evans and Honkapohja (2001).

The majority of this literature on learning in macroeconomics tends to focus on asymptotic results of convergence. In particular, an equilibrium is said to be E-stable (see Evans and Honkapohja, 2001) if it can represent the limiting point of the learning dynamics of agents. Little emphasis, instead, is usually put on the issue of how easy it is for agents to learn such an equilibrium. A notable exception in this regard is represented by Ferrero (2007), who, building on results from Marcet and Sargent (1995), considers the impact of monetary policy on the speed of convergence of the learning process of agents in an IS/AS economy. Berardi and Duffy (2010) also consider the issue of convergence times for different learning algorithms.

Recent literature has then focused on heterogeneity in learning and has tried to understand how heterogeneity affects the learnability of an equilibrium. See for example Giannitsarou (2003), Guse (2005), Branch and Evans (2006), and Berardi (2007). Also this literature, though, tends to focus on asymptotic results of convergence and neglects the issue of how heterogeneity affects the speed of convergence of the economy towards equilibrium.

The purpose of this paper is fill this gap and to assess the impact of heterogeneity on the speed of convergence of the learning dynamics of agents. We think that, though largely neglected, the question of the speed of convergence of an economy towards equilibrium is a very important one. It may help researchers, for example, decide whether, in a particular model, focusing only on

the limiting equilibrium point is a sensible thing to do, or if instead, due to the low speed of convergence, the economy will spend most of its time out of equilibrium and therefore transition dynamics should be considered carefully.

In a simple, forward looking linear stochastic model we thus investigate the impact of heterogeneity in expectations formation on the speed of convergence of the learning process of agents towards equilibrium. We do it using the Heterogeneous Expectations Equilibrium (HEE) concept developed in Berardi (2007), where we showed how heterogeneity affects learnability in terms of asymptotic results.

The structure of the paper is as follows: Section 2 introduces the model; Section 3 discusses E-stability and speed of convergence for the homogeneous case; Section 4 introduces heterogeneity and assesses its impact on E-stability and speed of convergence; Section 5 concludes.

2 The model

We borrow the model from Berardi (2007). The structural equations representing the dynamics for the economy are

$$x_t = aE_t^1 x_{t+1} + bE_t^2 x_{t+1} + cx_{t-1} + dw_t \quad (1)$$

$$w_t = \rho w_{t-1} + v_t, \quad (2)$$

where x_t is an endogenous variable, w_t is exogenous and observable at time t and v_t is an i.i.d. random disturbance. E^i indicates expectations of type i , not necessarily rational. The autoregressive parameter ρ is assumed to be between 0 and 1.

In this model agents can hold one of two types of expectations. Parameters a and b can be decomposed in two terms, one referring to the way in which expectations feed through on the endogenous variable and the other representing a weight on expectations that depends on the proportion of agents sharing the

same beliefs. Thus we could write

$$a = \alpha_1 \mu \tag{3}$$

$$b = \alpha_2(1 - \mu) \tag{4}$$

where μ is a measure of the relative size of the group of agents holding type 1 expectations. In particular, given a continuum of agents over the unit interval, μ of them are of type 1, and the remaining $(1 - \mu)$ are of type 2. Structural heterogeneity arises for $\alpha_1 \neq \alpha_2$. With structural homogeneity, $\alpha_1 = \alpha_2 = \alpha$, and if also expectational homogeneity holds (i.e., $E^1 = E^2 = E^C$), then $(aE^1 + bE^2) = \alpha E^C$, where E^C stands for common expectation, and we have the structural equation

$$x_t = \alpha E_t^C x_{t+1} + c x_{t-1} + d w_t \tag{5}$$

that substitutes (1).

3 Equilibrium, learnability and speed of convergence: the homogeneous case

It is well known (see, e.g., Berardi (2007)) that equations (1)-(2), under homogeneity, admit a minimum state variables rational expectations equilibrium (MSV REE) that, in its common factor representation (see Evans and McGough, 2005) takes the form¹

$$x_t = e_1 x_{t-1} + \frac{d}{1 - \alpha(e_1 + \rho)} w_t, \tag{6}$$

where

$$e_1 = \frac{1 - \sqrt{1 - 4\alpha c}}{2\alpha} \tag{7}$$

is the smaller root of matrix

$$A = \begin{pmatrix} \frac{1}{\alpha} & -\frac{c}{\alpha} \\ 1 & 0 \end{pmatrix}$$

¹For simplicity we consider in this work only learning of the common factor representation and leave for future work the analysis of learning and speed of convergence for the general form representation.

representing system (5) in first order form.²

If we then assume that agents do not have rational expectations from the start, but need to form their forecasts much like an econometrician does, and periodically update parameters in their forecasting model (or, in the learning literature terminology, their perceived law of motion - PLM) based on the new information becoming available at each time, we can consider the question whether (6) can represent the limiting point of an economy populated by such learning agents. In particular, assume that agents use the PLM³

$$x_t = \phi_1 x_{t-1} + \phi_2 w_t \tag{8}$$

and periodically update parameters ϕ_1 and ϕ_2 using recursive least squares techniques. It can then be shown (Berardi, 2007) that equilibrium (6) is learnable (E-stable) for

$$\alpha\rho < \frac{1 + \sqrt{1 - 4\alpha c}}{2}. \tag{9}$$

3.1 Speed of convergence

Convergence of the learning process is governed by a map from the PLM to the temporary equilibrium (or actual law of motion - ALM) that is obtained by inserting the estimated model (8) projected forward one period into the structural equation (5) in place of the expectations. In other words, when agents use a forecasting model consistent with solution (6) and update its parameters over time, at each point in time, for a given set of belief parameters, we have a temporary equilibrium (ALM) that represents the dynamics of the economy under a certain set of beliefs. A T-map can then be derived that maps parameters in the PLM (the beliefs of agents) into parameters of the ALM (the law of motion

²Note that, for the equilibrium to be determinate (i.e., locally unique), matrix A must have one stable and one unstable root. This happens for $\alpha < 1 - c$. In the remaining part of the paper, when the equilibrium is not determinate, we will focus on the MSV solution, which is always unique by constaction (see McCallum, 1983).

³We assume here that agents know the long run mean of the endogenous variables (here equal to zero), so they don't have to include an intercept in the PLM.

for the economy) as follows: starting from the forecasting model or PLM:

$$x_t = \phi_1 x_{t-1} + \phi_2 w_t$$

and inserting it into the structural model (5), we obtain the ALM:

$$x_t = \alpha(\phi_1(\phi_1 x_{t-1} + \phi_2 w_t) + \phi_2 \rho w_t) + c x_{t-1} + d w_t = (\alpha\phi_1^2 + c) x_{t-1} + (\alpha\phi_2(1 + \rho) + d) w_t.$$

Mapping term by term the PLM to the corresponding parameter in the ALM, we derive the T-maps:

$$\phi_1 \rightarrow \alpha\phi_1^2 + c \quad (10)$$

$$\phi_2 \rightarrow \alpha\phi_2(\phi_1 + \rho) + d. \quad (11)$$

The REE (6) is a fixed point $(\bar{\phi}_1, \bar{\phi}_2)$ of these T-maps, where

$$\bar{\phi}_1 = e_1 \quad (12)$$

$$\bar{\phi}_2 = \frac{d}{1 - \alpha(e_1 + \rho)}. \quad (13)$$

From T-maps (10)-(11), it is then possible to derive a system of ODEs that govern E-stability (see Evans and Honkapohja (2001) for a detailed explanation of the techniques):

$$\dot{\phi}_1 = \alpha\phi_1^2 + c - \phi_1 \quad (14)$$

$$\dot{\phi}_2 = \alpha\phi_2(\phi_1 + \rho) + d - \phi_2. \quad (15)$$

E-stability obtains if the system of ODEs is locally asymptotically stable at the equilibrium point $(\bar{\phi}_1, \bar{\phi}_2)$, i.e., if the eigenvalues of the Jacobian of the system, evaluated at $(\bar{\phi}_1, \bar{\phi}_2)$, have all negative real part.

The same ODEs can then be used to assess the speed of convergence of the learning process when implemented through recursive least squares (RLS) estimation. These ODEs in fact can be derived, using stochastic approximation techniques, from the stochastic recursive algorithm representing RLS in real time. Denoting by $\beta = (\phi_1, \phi_2)$ the vector of parameters to be estimated and

by z the vector of state variables, RLS is implemented by

$$\beta_t = \beta_{t-1} + t^{-1}R_t^{-1}Q(\beta_{t-1}, z_t). \quad (16)$$

If we then let

$$h(\beta) = E[Q(\beta, z)],$$

where the expectation is taken over the invariant distribution of z for given β , we have that $h(\beta)$ determines the T-maps from parameters in the PLM to those in the ALM and thus governs both the speed of convergence (see Benveniste et al, 1990) and E-stability (Evans and Honkapohja, 2001). These T-maps determine then the system of ODEs shown above.

Using theoretical results from Benveniste et al (1990), Theorem 3, p. 110, it is possible to show that root-t convergence of estimated parameters β obtains if the eigenvalues of the Jacobian of the system, evaluated at $\bar{\beta} = (\bar{\phi}_1, \bar{\phi}_2)$, all have real part smaller than $-1/2$.⁴ The set of parameters for which root-t convergence obtains is therefore a subset of the set of parameters for which E-stability obtains.

In our case, square root-t convergence obtains for $\alpha c < 3/16$ and $\alpha^2\rho^2 + \alpha c < 1/4$,⁵ where α is the parameter on the forward looking term and c the parameter on the backward looking term in the structural model (5), while ρ represents the degree of persistence of the exogenous driving variable w_t .

Proposition 1 *Parameters ϕ_1 and ϕ_2 in (8), when updated in real time by agents through the RLS algorithm (16), converge respectively to their equilibrium values $\bar{\phi}_1$ and $\bar{\phi}_2$ in (12) and (13) at root-t speed if $\alpha c < 3/16$ and $\alpha^2\rho^2 + \alpha c < 1/4$.*

For different values of α , c and ρ , we show the region where E-stability obtains (brown region in Fig. 1) and root-t convergence obtains (brown region

⁴Equivalently, the slope of the T-map has to be $< 1/2$.

⁵Moreover, to get a real solution, the following condition must be satisfied: $\alpha c < 1/4$. It can be easily verified that this condition is always satisfied if root-t convergence obtains.

in Fig 2): as expected from theoretical results, the region of root-t convergence is a subregion of the one of E-stability.

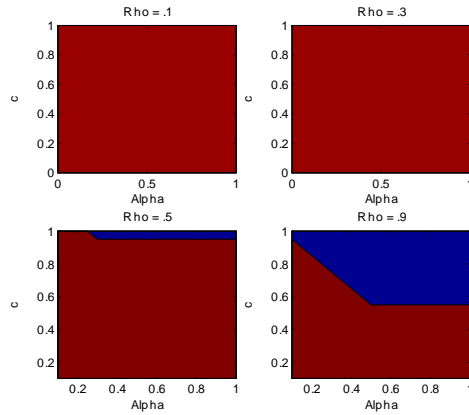


Figure 1: E-stability (brown region).

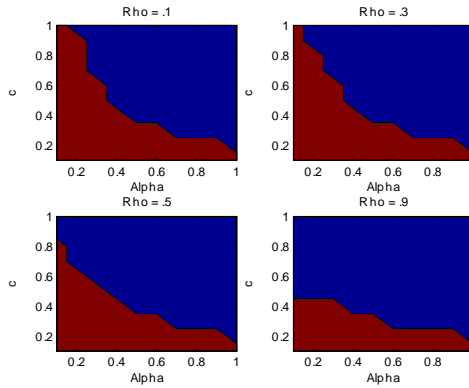


Figure 2: Root-t convergence (brown region).

Therefore, just as E-stability does, the requirement of root-t convergence poses a restriction on the structural parameters of the model: only if this restriction is satisfied the learning process of agents converges to its limiting equilibrium point at the speed at which, in classical econometrics, the mean of the distribution of the least squares estimates converges to the true value of the

estimated parameters.

By looking at the regions where root-t convergence obtains in Fig. 2, we can see that as ρ increases this region shrinks down, and the same happens as α and c increase. In general, therefore, the speed of convergence is higher the lower is the degree of persistence of the exogenous shock, the lower is the impact of expectations on current outcomes and the lower is the degree of intrinsic persistence in the endogenous variables. In other words, convergence is faster when there is little correlation across time in the structure of the economy.

4 Equilibrium, learnability and speed of convergence: the heterogeneous case

In the heterogeneous setting, we assume that there are two classes of agents, of size μ and $(1 - \mu)$, that form expectations using respectively the model PLM^1

$$x_t = \phi_1 x_{t-1} + \phi_2 w_t \quad (17)$$

and PLM^2

$$x_t = \theta w_t. \quad (18)$$

While the first model is consistent with the MSV REE, the second one is underparameterized, as it misses the lagged endogenous variable on the r.h.s..

The T-maps from PLM to ALM for the two groups of agents in this case are

$$\phi_1 \rightarrow a\phi_1^2 + c \quad (19)$$

$$\phi_2 \rightarrow a\phi_2(\phi_1 + \rho) + b\rho\theta + d \quad (20)$$

$$\theta \rightarrow \frac{a\phi_2(\phi_1 + \rho) + b\rho\theta + d}{1 - \rho(a\phi_1^2 + c)} \quad (21)$$

where the second is obtained by projecting the ALM onto the restricted space of variables in the PLM^2 (see Berardi (2007) for the derivation of these T-maps). Fixed point of these T-maps are the possible equilibria of this system. Berardi (2007) shows that an heterogeneous expectations equilibrium (HEE) loosely

defined as a triple $(\bar{\phi}_1, \bar{\phi}_2, \bar{\theta})$ that represents a fixed point for the system (19)-(21)⁶ exists. In particular, $\bar{\phi}_1$ is as in (12), while $(\bar{\phi}_2, \bar{\theta})$ are obtained numerically by solving the system (20)-(21) for a fixed point. Again, the slope of these T-maps at the HEE $(\bar{\phi}_1, \bar{\phi}_2, \bar{\theta})$ governs both E-stability of the equilibrium and the speed of convergence of the learning process of agents.

The equilibrium turns out to be learnable (E-stable) if all the eigenvalues of matrix

$$\Gamma = \begin{pmatrix} \frac{1-\sqrt{1-4ac}}{2} + a\rho - 1 & b\rho \\ \frac{\frac{1-\sqrt{1-4ac}}{2} + a\rho}{1-\rho(\frac{1-\sqrt{1-4ac}}{2a})} & \frac{b\rho}{1-\rho(\frac{1-\sqrt{1-4ac}}{2a})} - 1 \end{pmatrix} \quad (22)$$

have negative real part, which reduces to

$$\frac{1 - \sqrt{1 - 4ac}}{2} + a\rho + \frac{b\rho}{1 - \rho(\frac{1-\sqrt{1-4ac}}{2a})} < 1. \quad (23)$$

Figure 3 shows the E-stability region for different values of α (we assume for simplicity $\alpha_1 = \alpha_2 = \alpha$), c and μ :

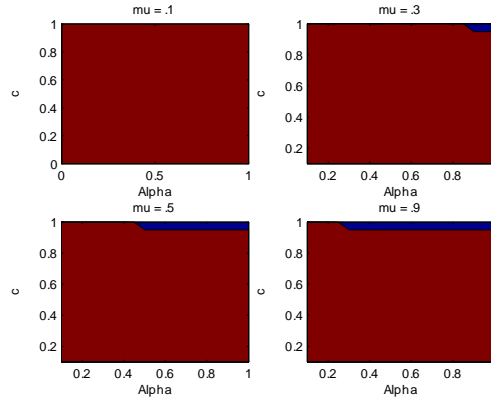


Figure 3: E-stability of HEE (brown region).

4.1 Speed of convergence

As pointed out above, and shown in detail in Berardi (2007), the fraction of agents (μ) using each of the two forecasting models available affects the learn-

⁶For a formal definition of the concept of HEE, we refer the reader to Berardi (2007).

ability (E-stability) properties of the heterogeneous equilibrium. The question we want to address here is how the same fraction of agents using each model affects the speed of convergence of the learning process, once learnability is established. In other words, we want to refine results from Berardi (2007) in order to understand the whole impact of heterogeneity on learning dynamics.

Analyzing the slope of the T-maps (19)-(21), it is easy to verify that root-t convergence obtains if $ac < 3/16$ and all eigenvalues of matrix Γ have real part $< -1/2$. The first condition implies that, as μ increases, root-t convergence is less likely to obtain.

Proposition 2 *Parameters ϕ_1 , ϕ_2 and θ in (17) and (18), when updated in real time by agents through a RLS algorithm equivalent to (16), converge respectively to their equilibrium values $\bar{\phi}_1$, $\bar{\phi}_2$, and $\bar{\theta}$ at root-t speed if $ac < 3/16$ and all eigenvalues of matrix Γ in (22) have real part smaller than $-1/2$.*

Figure 4 shows the region where root-t convergence obtains (brown area) for different values of α (we assume for simplicity $\alpha_1 = \alpha_2 = \alpha$), c and μ .⁷

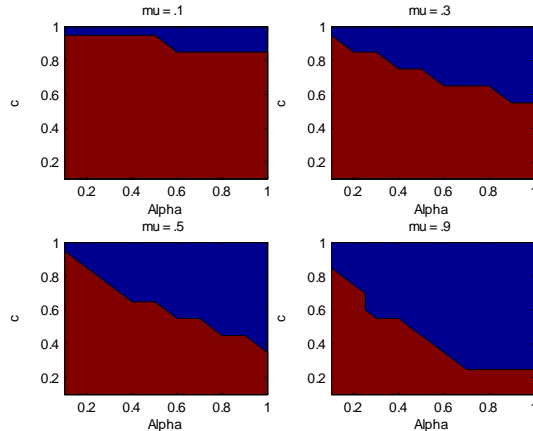


Figure 4: Root-t convergence for HEE (brown region).

Comparing Fig. 3 and Fig. 4, we can see that there are regions of the

⁷Here ρ is set to 0.5.

parameter space where convergence obtains (all eigenvalues λ_i of matrix Γ have negative real part), but such convergence is slower than root-t convergence (i.e., $\exists \lambda_i : -1/2 < \text{re}(\lambda_i) < 0$).

The main result here is that, as μ increases, even when the HEE remains E-stable the region for which root-t convergence obtains shrinks. This can also be seen by comparing Fig. 4 with Fig. 2 and noting that the third quadrant of Fig. 2 represents the limit point of the HEE represented in Fig. 4 as $\mu \rightarrow 1$: a higher fraction of agents in the economy using the forecasting model PLM^1 slows down convergence of the learning process.

In order to assess the speed of convergence of learning when slower than root-t, we follow Marcat and Sargent (1995) who propose a numerical procedure to estimate the rate of convergence when E-stability obtains (i.e., the procedure assumes that the learning process converges at a certain speed, and estimates such a speed).

In particular, we define the rate of convergence δ as

$$t^\delta (\beta_t - \bar{\beta}) \xrightarrow{d} F \quad (24)$$

where β_t is the estimated parameter, $\bar{\beta}$ is its equilibrium value and F is a non degenerate distribution with mean zero and variance σ_F^2 . We can then obtain a numerical approximation of δ for large t : since (24) implies

$$\lim_{t \rightarrow \infty} E [t^\delta (\beta_t - \bar{\beta})]^2 = \sigma_F^2$$

we have

$$\frac{E [t^\delta (\beta_t - \bar{\beta})]^2}{E [(tk)^\delta (\beta_t - \bar{\beta})]^2} \rightarrow 1$$

and therefore

$$\frac{E(\beta_t - \bar{\beta})^2}{E(\beta_{tk} - \bar{\beta})^2} \rightarrow k^{2\delta} \text{ as } t \rightarrow \infty.$$

This implies that we can estimate δ with

$$\delta = \frac{2}{2 \log(k)} \log \left(\frac{E(\beta_t - \bar{\beta})^2}{E(\beta_{tk} - \bar{\beta})^2} \right)$$

for large t . This can be done practically by choosing t and k and approximate the expectational terms by Monte Carlo integration, i.e., by simulating a large number of series of length t and tk and calculating the mean square error across realizations.

We thus compute numerically δ by computing the mean square error across 1000 realizations for $t = 3000$ and $tk = 15.000$. We set our baseline parameterization as follows: $\alpha_1 = \alpha_2 = c = d = \rho = .5$.⁸ Numerical results are reported in Table 1.

	δ_{ϕ_1}	δ_{ϕ_2}	δ_{θ}
$\mu = .1$.5164	.4564	.4820
$\mu = .5$.5212	.4145	.4900
$\mu = .9$.3018	.2652	.3634

Table 1: Speed of convergence for parameters in PLMs: baseline parameterization

By looking at Fig. 4, we would expect, for $\mu = .1$ and $.5$, $\delta \simeq .5$, while for $\mu = .9$, we should observe $\delta < .5$. In fact, we can see a sharp drop in the speed of convergence as we move from the second to the third row in Table 1. We must also remember though that, as noted by Marcet and Sargent (1992), only for rather large t and tk the asymptotic results of the theorem by Benveniste et al. hold, while for smaller samples numerical results can deviate largely from the theoretical ones. In particular, in our case, we observe that the learning speed of parameter ϕ_2 in PLM^1 is constantly below root- t for finite sample sizes.

Our simulations also confirm that, as μ increases, the speed of convergence decreases below root- t . These results mean that even when heterogeneity does not affect E-stability (in all cases reported in Table 1, the HEE is E-stable), it still affects the speed of convergence. Interestingly, the higher is the fraction of agents using the simpler, underparameterized, model, the faster is convergence.

Remark 3 *An increase in μ , by increasing a , moves the system towards a*

⁸We conducted extensive robustness checks for different parameterizations and the main results reported here hold throughout.

region where convergence of the learning process of agents is slower than root-t.

Remark 4 *When E-stability prevails but convergence is slower than root-t, an increase in μ decreases the speed of convergence for the learning process of agents.*

Table 2 reports results for two alternative parameterizations for α and c (while all other parameters have been kept fixed). In this case root-t convergence obtains for all cases reported, and we can see that in fact μ does not affect much the estimated speed of convergence.

		δ_{ϕ_1}	δ_{ϕ_2}	δ_{θ}
$\alpha = .5, c = .1$				
$\mu = .1$.5008	.4730	.4704
$\mu = .5$.5161	.5031	.5092
$\mu = .9$.5003	.4790	.4794
$\alpha = .1, c = .5$				
$\mu = .1$.5175	.4880	.4924
$\mu = .5$.5088	.5189	.5073
$\mu = .9$.4829	.4794	.5005

Table 2: Speed of convergence for parameters in PLMs: robustness ckeck

Note that μ is the fraction of agents using a correctly specified forecasting model. For $\mu = 1$, the economy is back in the homogeneous case of section 3, where in fact we saw that, for the baseline parameterization, convergence is slower than root-t. For $\mu = 0$, instead, all agents use a misspecified model and the economy can only converge to a restricted perceptions equilibrium (RPE).⁹ This RPE turns out to be E-stable for

$$\frac{b\rho}{1 - \rho c} < 1 \quad (25)$$

and the learning process of agents converges to this RPE at root-t for

$$\frac{b\rho}{1 - \rho c} < \frac{1}{2}. \quad (26)$$

⁹For a formal definition of a RPE, see Evans and Honkapohja (2001).

In the case of our baseline parameterization, with $\alpha_2 = \rho = c = .5$, this condition is in fact satisfied (the l.h.s. of (26) is equal to $1/3$), and the learning process of agents in group 2 converges at root-t speed.

In other words, if our economy is populated only by agents using a correctly specified forecasting model, it will converge towards the REE, but at a speed slower than root-t, while if our economy is populated only by agents using a misspecified (underparameterized) model, it will converge towards a RPE equilibrium, but at the faster, root-t speed.

There seems to be therefore a trade-off between speed of convergence and quality of forecasts in equilibrium: the larger is the fraction of agents using the correctly specified model, the slower is convergence, but the closer is the equilibrium thus emerging to the REE.¹⁰ On the other end, when a high fraction of agents uses an underparameterized model, convergence is faster, but towards an equilibrium that resembles more and more a RPE. In a RPE agents, though forming their expectations efficiently given their misspecified model, are not rational, and have larger forecast errors.

It is also interesting to compare the speed of convergence of the learning process across the two groups of agents. While agents in group 1 have to learn two parameters, one for the lagged endogenous variable x_{t-1} and one for the exogenous driving process w_t , agents in group 2 only learn one parameter, for the exogenous variable. By comparing the speed of convergence for the learning processes of the two groups for w_t , we can see that while when root-t convergence obtains the learning speed for the two groups of agents is similar, when root-t convergence does not obtain the learning for agents in group 2 is always faster to converge: this means that neglecting the lagged endogenous variable in the algorithm speeds up the learning process for the parameter on the (only) other variable.

Remark 5 *When the learning process of agents converges at a rate slower than*

¹⁰By "closer" here we mean that there is a higher fraction of agents having expectations that can be considered rational.

root- t , agents using an underparameterized PLM learn the equilibrium value for the parameters in their model more quickly than agents using a correctly specified model.

5 Conclusions

We have presented in this work a simple model of heterogeneous expectations and learning, where agents use different models to form their expectations, and we have analyzed the impact of the degree of heterogeneity on the speed of convergence for the learning process. We have found that the proportion of agents using each forecasting model affects the speed of convergence, and that the larger is the number of agents using the simpler, but misspecified, model, the faster is convergence towards equilibrium. There seems to be a trade-off, therefore, between the accuracy of the expectations in the population and the speed by which agents can learn the equilibrium. Moreover, agents with a misspecified (underparameterized) model learn faster the equilibrium value for their beliefs.

We think these results are important and require attention from scholars interested in using learning as a coordination device that can lead the economy towards an equilibrium: as shown here, the speed of convergence of the learning activity is in general dependent on the various parameters of the model, a crucial one being the degree of heterogeneity in the economy.

References

- [1] Benveniste, A., Metivier, M., Priouret, P., 1990. Adaptive Algorithms and Stochastic Approximations. Springer, Berlin.
- [2] Berardi, M., 2007. Heterogeneity and misspecifications in learning. *Journal of Economic Dynamics and Control* 31(10), 3203-3227.
- [3] Berardi, M., Duffy, J., 2010. Real time adaptive learning via parameterized expectations. Centre for Growth and Business Cycle Research Discussion Paper Series, University of Manchester, n. 147.
- [4] Branch, W.A., Evans, G.W., 2006. Intrinsic heterogeneity in expectations formation. *Journal of Economic Theory* 127, 264-295.
- [5] Evans, G.W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.
- [6] Ferrero, G., 2007. Monetary policy, learning and the speed of convergence. *Journal of Economic Dynamics and Control* 31, 3006-3041.
- [7] Giannitsarou, C., 2003. Heterogeneous learning. *Review of Economic Dynamics* 6, 885-906.
- [8] Guse, E. A., 2005. Stability properties for learning with heterogeneous expectations and multiple equilibria. *Journal of Economic Dynamics and Control* 29, 1623-1642.
- [9] Marcet, A., Sargent, T.J., 1989. Convergence of least squares learning mechanism in self-referential linear stochastic models. *Journal of Economic Theory* 48, 337-368.
- [10] Marcet, A., Sargent, T.J., 1995. Speed of Convergence of Recursive Least Squares Learning with ARMA perceptions. In: Kirman, A., Salmon, M. (Eds.), *Learning and Rationality in Economics*, Basil Blackwell, Oxford, 179-215.

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