Are consistent expectations better than rational expectations ?

Elliot Aurissergues

Are consistent expectations better than rational expectations ?

Elliot Aurissergues

Abstract

In this paper, I argue that agents may prefer learning a misspecified model instead of learning the rational expectation model. I consider an economy with two types of agent. Fundamentalists learn a model where endogenous variables depend on relevant exogenous variables whereas followers learn a model where endogenous variables are function of their lagged values. A Fundamentalist is like a DSGE econometrician and a follower is like a VAR econometrician. If followers (resp. fundamentalists) give more accurate forecasts, a fraction of fundamentalists (resp. followers) switch to the follower model. I apply this algorithm in a linear model. Results are mixed for rational expectations. Followers may dominate in the long run when there are strategic complementarities and high persistence of exogenous variables. When additionnal issues are introduced, like structural breaks or unobservable exogenous variable, followers can have a significant edge on fundamentalists. I apply the algortihm in three economic models a cobweb model, an asset price model and a simple macroeconomic model.

JEL Classification: D83,D84

Keyword: Adaptive learning, cobweb model, naive expectations

Elliot Aurissergues

PHD student, Paris School of Economics and University Paris Pantheon Sorbonne

elliot.aurissergues@gmail.com

Introduction

In general equilibrium macroeconomic models, endogenous variables (prices and quantities) are determined by a system of supply and demand. These aggregate supply and demand are the result of individual policy rule. To choose their optimal policy, agents have to forecast future values of endogenous variables (or to give a probability distribution). If they know the relation between endogenous variables and exogenous variables, they are able to form their forecast by using available information about exogenous variables. Their policy rule will only depends on exogenous variables, validating agents belief. This is the rational expectation hypothesis. But what happens if agents do not know the parametric relation between endogenous variable and exogenous ones ? The learning literature have been developed to adress this question. Agents try to learn the fundamental model of the economy. More precisely, they estimate through econometric techniques the relation between endogenous and exogenous variables. It is well established that recursive learning converges to the rational expectation solution if strategic complementarities are not too "strong", or more precisely if the elasticity of the variable with repsect to its expectation is inferior to 1 (Bray and Savin, 1986). This result have strengthened the case for the REH, suggesting that the level of knowledge it requires is not so implausible.

My paper questions this conclusion. I do not challenge the convergence result of the learning literature. In simple terms, I ask if this learning strategy is really a good idea for economic agents. In particular, I compare it with an alternative strategy where agents learn the correlation between current endogenous variable and past ones. They are VAR econometricians instead of DSGE econometrician.

It is not obvious that rational expectation learning delivers more accurate forecasts than this pragmatic behavior. When agents learn, they make expectationnal errors which are correlated across time. By definition, exogenous variables do not deliver any information about these expectationnal errors. Past equilibrium outcomes do and they incorporate important information about exogenous variable if they are persistent. An agent can make a more accurate forecast by learning a backward looking model , even if she knows this is a misspecified relation.

To give a concrete example, think about a DSGE model. To choose its consumption level, the representative consumer should form expectations about all future values of real wages. In the REH model, These values are determined by "shocks", for example productivity and monetary policy shocks. If agents do not know the relation between real wages and shocks, they may learn it. This is the fundamental learning. The alternative strategy is to infer future wages from past wages. Past wages have the advantage to incorporate information about unobservable variable like expectations of other agents, unobserved shocks or structural breaks.

To decide between the two learning schemes is difficult in an environment with limited knowledge. An agent cannot choose optimally between the two. It would imply to make an optimal forecast of the forecasting error of the two models which depends on the model which is chosen. To adress the issue, I adopt an evolutionnary viewpoint. I see the adoption of different models as a process of trials and errors in the spirit of Nelson and Winter (1982), or more recently Saint Paul (2015). I simulate an economy with two type of forecasters, fundamental learners and adaptative learners. At each period, forecasts delivered by the two strategies are compared with the actual outcome. A fraction of the agents using the less accurate strategy shifts to the other one. I look at the convergence toward one strategy. I repeat this exercise in three models, a cobweb model, an asset price model, and a simple kyenesian macroeconomic model.

1 Framework

Consider a model in which a macroeconomic variable x is given by the equation

$$x_t = \alpha + \beta y_t + \lambda x_t^E + u_t \tag{1}$$

α is a constant.β, λ are parameters.u_t is a white noise of standard deviation σ_u. It is not observable by agents before they make their decision in period t.

y is an exogenous variable which follows an autoregressive process

$$y_t = \theta y_{t-1} + \epsilon_t \tag{2}$$

θ is a parameter and ϵ_t is a white noise of standard deviation σ_y

x_t^E is the expectation of the value of x in t. There is a mass one of agents. Each of them form a forecast.x^E is the aggregate forecast.

$$x_t^E = \int_0^1 x_t^e(i)di \tag{3}$$

$x_t^e(i)$ is the indiviudual expectation formed at period t by agent i.

ζ_t is a white noise.

Agents observe contemporaneous values of exogenous variables and lagged endogenous variables. Formally, their information set can be summarized by

$$\Theta_t = \{y_T, x_{T-1}\}_{T=-\infty}^{T=t}$$

Rational expectations solution It is convenient to compute the rational expectation solution of the model before outlining the main point.

The fixed point value for x, denoted x^* is given by

$$x^* = \frac{\alpha}{1-\lambda}$$

I denote $\hat{x_t} = x_t - x^*$. Equation (1) can be rewritten

$$\hat{x_t} = \beta y_t + \lambda \hat{x}_{E,t} + u_t$$

The rational expectation value for $\hat{x}_t$ can be found easily. One way is to use undetermined coefficient method. I guess that the deviation of x from steady state depends from the last observation of the exogenous variable y_{t-1}.

$$\hat{x_t} = \frac{\beta}{1-\lambda} y_t + u_t$$

The learning hypothesis The rational expectation for x_t is

$$E(x_t) = \frac{\alpha}{1-\lambda} + \frac{\beta}{1-\lambda} y_t$$

To form the rational expectation, the agent have to know the values of $\frac{\alpha}{1-\lambda}$ and $\frac{\beta}{1-\lambda}$. What happens if they do not know them ? The learning literature was developed to answer this question.

An agent will try to learn the values of the two parameters. In period t, She estimates the model

$$x_k = \pi_f y_k + \varphi_f + u_{f,k}$$

where k goes from period 0 (where observables start) to period $t-1$.

After estimating the model, they form a forecast for x_t

$$x_t^e = \pi_{f,t} y_t + \varphi_{f,t}$$

I label these agents "fundamentalists".

It has been shown by Bray and Savin (1986) that the estimator $(\hat{\pi}, \hat{\varphi})$ converges toward $\frac{\alpha}{1-\lambda}, \frac{\beta}{1-\lambda}$ if $\lambda < 1$

The alternative forecast This learning strategy converges toward the "good" solution (according to economic theory). But, does it allows agents to make accurate forecasts ?

I explore the possibility that an alternative strategy provides better forecast. where agents do not learn the "true" model. If they do, agents will adopt them, and the rational expectation solution is misspecified.

In particular, I consider a strategy where agents learns autocorrelation for endogenous variables. I label this strategy the "follower" strategy. Followers learn the model

$$x_k = \pi_o x_{k-1} + \varphi_o + u_{o,t}$$

In period t, their forecast for x is

$$x_t^e = \pi_{o,t} x_{t-1} + \varphi_{o,t}$$

Followers behave like VAR econometrician whereas fundamentalists can be viewed as "DSGE" econometrician. Intuitively, it seems difficult to believe that the follower strategy could deliver more accurate forecasts than the fundamentalist one. Indeed, the follower strategy does not use all availabe information. In particular, it does not take into account contemporaneous innovation on y whereas it takes into account past innovation on x u_{t-1} which should not be relevant to forecast x_t. This reasoning is true if agents are fully informed about parameter values. But it is more complicated if they should first learn these values for two reasons. First, the fundamental learning is actually misspecified as aknowledged by the literature (Bray and Savin 1986), (Evans and Honkapoja 2004) and this misspecification can give advantages to followers. Second, if there are strategic complementarities, the follower strategy can be self fulfiling.

2 Multiple Equilibria

In this section, I study the long run behavior of the economy. I define long run as a situation in which both type of agents have a stable estimation of their respective models and in which there is only one type of agent remaining.

There is a long run equilibrium if the dominant agent makes more accurate forecasts in average than the other type. There are no incentives for the dominant agent to deviate from her model.

Two resuts emerge. First, the situation in whihc fundamentalists dominate is always an equilibrium. I label it the "fundamentalist equilibrium". Indeed, if there are only fundamentalists, their model is correctly specified in the long run and their average forecasts errors are equal to standard deviation of the white noise u. Followers have a misspecified model and make forecasts errors in average.

A second result is that situations in which followers dominate are also an equilbrium for a large set of parameters. I detail these results in following paragraphs.

Equilbrium definition First I define the two type of equilbriums. In the following definition and propositions, the limit of a sequence of random variables is a random variable toward which the sequence converges in probability. I define the two matrix of observables $Z_{o,T}$ and $Z_{f,T}$. These matrix have T columns and these columns are observations respectively for vectors $\begin{pmatrix} 1 \\ x_{t-1} \end{pmatrix} (1 \quad x_{t-1})$ and $\begin{pmatrix} 1 \\ y_t \end{pmatrix} (1 \quad x_{t-1})$. I also defines the vector X_T which is the vector column for observations of x from 0 to T

Definition 1 *A fundamentalist equilbrium is a couple of vectors* (π_o, φ_o), (π_f, φ_f) *for which*

1. $(\pi_o, \varphi_o) = \lim_{T\to\infty} \frac{1}{T}(Z'_{o,T} Z_{o,T})^{-1}(Z'_{o,T} X_T)$
2. $(\pi_f, \varphi_f) = \lim_{T\to\infty} \frac{1}{T}(Z'_{f,T} Z_{f,T})^{-1}(Z'_{f,T} X_T)$
3. $\forall T \quad x_T = \alpha + \lambda\varphi_f + (\beta + \lambda\pi_f) y_T + u_T$
4. $\lim_{T\to\infty} E[(x_T - \varphi_f - \pi_f y_T)^2] < \lim_{T\to\infty} E[(x_T - \varphi_o - \pi_o y_T)^2]$

The follower equilbrium is defined similarly

Definition 2 *A follower equilbrium is a couple of vectors* (π_o, φ_o), (π_f, φ_f)*, both belonging to* $\mathbb{R}^2$ *for which*

1. $(\pi_o, \varphi_o) = \lim_{T\to\infty} \frac{1}{T}(Z'_{o,T} Z_{o,T})^{-1}(Z'_{o,T} X_T)$
2. $(\pi_f, \varphi_f) = \lim_{T\to\infty} \frac{1}{T}(Z'_{f,T} Z_{f,T})^{-1}(Z'_{f,T} X_T)$
3. $\forall t \quad x_T = \alpha + \lambda\varphi_o + \beta y_T + \lambda\pi_o) x_{T-1} + u_T$
4. $\lim_{T\to\infty} E[(x_T - \varphi_f - \pi_f y_T)^2] > \lim_{T\to\infty} E[(x_T - \varphi_o - \pi_o y_T)^2]$

In a nutshell, at the equilibrium, algorithm have converged, there is only one type remaining and this type makes more accurate forecasts in average.

Here, I used square errors instead of absolute deviation errors which do not allow analytical results.

2.1 Fundamentalist Equilibrium

Fundamentalist forecasts and errors The equilibrium equation is given by

$$x_t = \alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)y_t + u_t$$

If λ is inferior to 1. it is known that

$$\pi_f = \frac{\beta}{1-\lambda}$$
$$\varphi_f = \frac{\alpha}{1-\lambda}$$

Errors are straightforward to compute

$$lim_{t\to\infty}E[(x_t - \varphi_f - \pi_f y_t)^2] = \sigma_u^2$$

Follower forecast and errors Followers estimation converges toward

$$\pi_o = lim_{t->\infty}\frac{cov(x_t, x_{t-1})}{V(x_t)}$$
$$\varphi_o = \frac{\alpha(1-\pi_o)}{1-\lambda}$$

Computations of covariance and variance gives

$$\pi_o = \theta - \frac{\sigma_u^2\theta}{\left(\frac{\beta}{1-\lambda}\right)^2\frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2}$$

If σ_u is small, θ is a good approximation of π_o. Using this approximaation, forecasts errors of followers are

$$lim_{t\to\infty}E[(x_t - \varphi_o - \pi_o y_t)^2] = (1+\theta^2)\sigma_u^2 + \left(\frac{\beta}{1-\lambda}\right)^2\sigma_y^2$$

Obviously, their squared errors are always bigger than fundamentalists squared errors. It leads to the following proposition

Proposition 1 *For all vector of parameters $(\beta, \alpha, \sigma_u, \sigma_y, \theta)$, there exist a fundamentalist equilibrium*

Whereas fundamentalists make more accurate forecasts in this situation. Squared errors of followers are not necessarily high compared to the variance of X. Indeed, variance of X is

$$V(x) = \left(\frac{\beta}{1-\lambda}\right)^2\frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2 \tag{4}$$

Suppose that σ_u is small compare to σ_y and that θ is close to one, the variance of x is several times the average squared errors of the autoregressive model. Intuitively, the model is misspecified but still explains a large part of x variations.

2.2 Follower equilibrium

Follower forecasts and errors If followers dominates, the equilibrium equation is given by

$$x_T = \alpha + \lambda\varphi_o + \beta y_T + \lambda\pi_o x_{T-1} + u_T \tag{5}$$

The coefficient of the autoregressive process is given by

$$\pi_o = \lim_{t->\infty} \frac{Cov(x_t, x_{t-1})}{V(x_t)} \tag{6}$$

However, both the covariance and the variance in the previous formula depends on π_o. the coefficient of the autoregressive process is given by the roots of the equation

$$\pi_o = \frac{\beta^2 V(y)(\theta + \lambda\pi_o) + \lambda\pi_o(1 - \lambda\pi_o\theta)\sigma_u^2}{\beta^2 V(y)(1 + \lambda\pi_o\theta) + (1 - \lambda\pi_o\theta)\sigma_u^2} \tag{7}$$

It is a quadratic equation and thus an explicit solution for the two roots is possible. However, I choose to present a simpler case. Indeed, if $\sigma_u^2 = 0$, π_o is given by the polynomial equation

$$\pi_o^2\lambda\theta + \pi_o(1 - \lambda) - \theta = 0 \tag{8}$$

No simulations I made for this model converge toward the negative root. I choose to focus on the positive root. The positive root of the equation

$$\pi_o = \frac{\sqrt{(1-\lambda)^2 + 4\lambda\theta^2} - (1 - \lambda)}{2\lambda\theta} \tag{9}$$

Squared errors of followers are equal to

$$lim_{t\to\infty} E[(x_t - \varphi_o - \pi_o y_t)^2] = \beta^2 V(y) + (1-\lambda)^2\pi_o^2 V(x) + \sigma_u^2 + (1-\lambda)\pi_o\beta\theta cov(x, y) \tag{10}$$

3 The simulated economy

The fundamentalist algortithm Fundamentalists believe that the variable x can be forecasted by estimating the equation

$$x_t = \pi_f y_t + \varphi_f + u_{f,t} \tag{11}$$

I shortly describe the recursive algorithm used by fundamentalists to estimate (7).

I define the vector of exogenous variable $z_{f,t}$ and the vector of estimated parameters $\Phi_{f,t}$

$$z_{f,t} \equiv (1 \quad y_t)'$$
$$\Phi_{f,t} \equiv (\varphi_{f,t} \quad \pi_{f,t})'$$

I denote the variance covariance matrix $R_{f,t}$. Parameters estimates are updated by the two recursive equations.

$$R_{f,t+1} = R_{f,t} + \frac{1}{t}(z_{f,t}z_{f,t}^{'} - R_{f,t}) \tag{12a}$$

$$\Phi_{f,t+1} = \Phi_{f,t} + R_{f,t+1}\frac{1}{t}z_t\left(x_t - z_t^{'}\Phi_{f,t}\right) \tag{12b}$$

At period t, the forecast of fundamentalist is

$$x_{f,t} = \pi_{f,t}y_t + \varphi_{f,t} \tag{13}$$

The Follower algorithm Followers have a different strategy. They believe that the variable x is given by

$$x_t = \pi_o x_{t-1} + \varphi_o + \sigma_t \tag{14}$$

Like fundamentalists, they try to learn the value of π_o and the value of φ_o.

I introduce the vector of exogenous variable $z_{o,t}$ and the vector of estimated parameters $\Phi_{o,t}$

$$z_{o,t} \equiv (1 \quad x_{t-1})'$$

$$\Phi_{f,t} \equiv (\varphi_{o,t} \quad \pi_{o,t})'$$

The variance covariance matrix is $R_{o,t}$.

The recursive estimation is given by

$$R_{o,t+1} = R_{o,t} + \frac{1}{t}(z_{o,t}z_{o,t}^{'} - R_{foll,t}) \tag{15a}$$

$$\Phi_{o,t+1} = \Phi_{o,t} + R_{o,t+1}\frac{1}{t}z_{o,t}\left(x_t - z_{o,t}^{'}\Phi_t\right) \tag{15b}$$

At period t, the forecast of followers is

$$x_{o,t} = \pi_{o,t}x_{t-1} + \varphi_{o,t} \tag{16}$$

Update of the share of followers Initially, there are $1-\gamma$ fundamentalists and γ followers.

At the end of period t, agents observe forecasts of both types $x_{f,t}, x_{o,t}$ and the actual outcome x_t.

After t simulated periods, they compute the statistics[1]

$$\Delta_t = \frac{1}{t}\sum_{n=0}^{t}\sqrt{(x_{o,t} - x_t)^2} - \frac{1}{t}\sum_{n=0}^{t}\sqrt{(x_{f,t} - x_t)^2} \tag{17}$$

[1] I also consider the alternative statistics $\Delta_t = \frac{\sum_{n=0}^{t}\mathbb{1}_{\{(\tilde{x}_{o,t}-x_t)^2<(\tilde{x}_{f,t}-x_t)^2\}}}{t} - 0.5$. Which is simply the number of times for which follower strategy have delivered a more accurate forecasts than the fundmaentalist strategy

The statistics Δ is simply the average forecasting error of the follower strategy minus the average forecasting error of the fundamentalist strategy

If $\Delta < 0$, the follower strategy was in average more accurate than the fundamentalist one until the period t, a fraction μ of the fundamentalists shifts to the follower strategy. Conversely, if the fundamentalist strategy have been more accurate in average, the same fraction shifts from the follower strategy to the fundamentalist one

Thus, the evolution of γ is given by

$$\gamma_{t+1} = \gamma_t - \mu\gamma_t \mathbb{1}_{\{\Delta_t > 0\}} + \mu(1 - \gamma_t)\mathbb{1}_{\{\Delta_t < 0\}} \tag{18}$$

Summary of the evolution of the model The structure of the model can be summarized by nine equations.

A first bloc of equations is composed of equilibrium equations. It includes the two forecast equation (9) and (12) and the equation giving the equilbrium value of x

$$x_t = \alpha + \beta y_t + \lambda\gamma_t x_{o,t} + \lambda(1 - \gamma_t)x_{f,t} + \zeta_t \tag{19}$$

There are two dynamic blocs.

The first dynamic bloc includes the two recursive estimation algorithm (8) and (11). which gives four equations.

A second dynamic bloc gives the evolution of the share of followers. These are thequations (13) and (14)

Description of the economy algorithm I simulate this economy over a long period. The algorithm may be summarized by the following sequence of events

1. Using the model they have chosen, their past estimates of parameter values and the value of y_t, x_{t-1}, fundamentalists and followers compute their forecasts for x_t

2. The equilibrium value of x_t

3. This value is compared with forecasts of both type of agents.

4. If the model of fundamentalists underperforms (respectively followers) model underperforms, they switch to the other model with probability μ.

5. Once they have chosen their new model, they estimate it using the history of values for x and y

Initialization The initial value for the share of fundamentalists is γ_0. A more significant practical issue is the initialization of the two learning algorithm. For both, there are four parameters to initialize, the covariance, the variance of the regressor, the constant and the coefficient of the regressor and the outcome of the simulation is sensitive to the choice of initialization. I look at a neutral initialization which does not give a large advantage to either fundamentalists or followers. My strategy is the following. I use long run values of parameters in the rational expectations equilibrium for both follower algorithm and fundamentalist one. I multiply these values by a constant ν.

Intuitively, ν represents how far priors are from the rational expectations equilibrium. A ν equal to 1 means that priors are set at the RE equilbrium and thus at the fundamentalist equilibrium.

4 Results

I simulated the economy for my baseline linear model, for the cobweb model and for the asset price model. Results for the linear and for the cobweb model are consistent. preliminary results for the asset price model are in line with the two others but the code need to be refined. the code for the macroeconomic model have yet to be build.

I start my simulation with an equal number of fundamentalists and followers $\gamma_0 = 0.5$. The initial value of γ does not seem affecting the long run behavior of the economy.

4.1 Main result

The main result of my simulations is that the economy does not always converge to the Rational Expextations Equilibrium. I illustrate the result by the figure (2). I simulate the economy for three different calibration of parameters λ and ν. In the first calibration, λ is set at 0.25 and ν at 0.9. In the second calibration, respective values are 0.65 and 0.9 and in the third one, λ is still equal to 0.65 but ν is set at 0.7. Other parameters are set in the following way. σ_y is equal to 0.01, σ_u to 0.005, θ to 0.8, and the initial share of fundamentalists $\gamma_0$is 0.5. In figure (2), I display the evolution of the share of fundamentalists for the three simulations. For the first calibration, The economy converges unambiguosly to the fundamentalist equilibrium. It does not occur for the two others. The share of fundamentalists rises in the beginning but then decreases toward zero. In figure (3) and (4), I display the evolution of forecasts errors for fundamentalists and followers for respectively the first and the third calibration.

4.2 Sensitivity

Six parameters matter for the ouctome. Four, $(\lambda, \theta, \sigma_y, \sigma_u)$ are related to the model itself and two $\nu, \gamma_0$are related to the initialization. In this paragraph, I explore how changes in parameter values affect the outcome of the simulation. Formally, I define the function $G : (\nu, \gamma_0, \lambda, \theta, \sigma_y, \sigma_u) \to \{0, 1\}$. The value of G is 0 (resp. 1 when the simulation run with parameter values $(\nu, \gamma_0, \lambda, \theta, \sigma_y, \sigma_u)$ converges to the follower equilibrium (resp. the fundamentalist equilibrium).

Representing a 7 dimension figure is difficult. I use my baseline calibration for four parameters and make two others varying. I consider succesively variations of (λ, ν), (θ, γ_0) and (σ_u, σ_y). Results are displayed in figure (5) to (7). The surface is in red when the simulation converges to fundamentalist equilibria and in blue if fundamentalists dominates. They confirm that a follower equilibrium is not a curiosity. The convergence toward the follower model occurs for a large set of parameters. The elasticity of the variable to expectations λ and the spread between initial priors and rational expectation values ν seems the more important determinant of the convergence.

4.3 Additionnal results

Some other results deserves to be highlighted.

Convergence toward RE equilibrium could be very slow when it occurs For some calibrations, the economy converges toward the fundamentalist equilbrium but very slowly. It may occur after several hundred periods of follower dominance. Modern macroeconomic databases contains at best 300 data points for one variable and datas are often characterized by a very high persistence and strong indications of structural breaks. It cast doubt on the practical relevance.

The role of parameter ν When ν is initialized to one, the economy seems always converging to the fundamentalist equilibrium. This equilibrium seems locally stable (when priors of algorithm are initially correct for this equilbrim) but not globally.

Moreover, I considered a uniform gap ν for all variables which have to be initialized. The goal was to limit the number of dimensions for the parameter space. However, I perform additionnal simulations in which the gap is different. More precisely, I consider different initial errors for the constant term and the autoregressive coefficient. These simulations suggests that the initial errors for the constant is the more important one.Intuitvely, fundamentalist take time to estimate accurately the constant. Followers are more efficient because their exogenous variable x_{t-1} also incorporates information about the constant.

The role of the parameter λ, comparison with e-stability

5 Extensions: misspecification and structural breaks

5.1 Intuition

Learning an autoregressive process for endogenous variables may be better than learning the true Rational Expectations model when agents know the "true" model of the economy. The advantage of the adaptative behavior could be even bigger if agents have a misspecified model of the economy. For example, the variable x may be affected by unobserved variables or the parameters of the equation (1) can be subject to structural breaks. Intuitively, past values of x can carry information about structural breaks or unobersved variables and an adapative behavior can capture it whereas purely exogenosu variables do not carry anything.

5.2 Adding persistent unobservables

A first misspecification is the existence of an unobserved exogenous variable. For example, the equation (1) becomes

$$x_t = \alpha + \beta y_t + \lambda x_t^E + u_t + v_t \tag{20}$$

with

$$v_t = \rho_v v_{t-1} + \epsilon_t^v$$

v is not observable by agents and fundamentalists continue to estimate the model $x_t = \varphi_f + \pi_f y_t$. Because v is persistent, past values of x carries information about the current value of v, giving an edge to followers over fundamentalists.

5.3 Adding structural breaks on constant

a second misspecification is that some parameters are not constant but time varying and follow for example a markov chain.

The equation becomes

$$x_t = \alpha_t + \beta y_t + \lambda x_t^E + u_t + v_t \tag{21}$$

α_t is a random variable whose support is the vector $\{\alpha_l, \alpha_h\}$, where both α_l and α_h are real numbers.

α_t evolves according to a markov chain. In the "h" state, the probability to remain in the high state is p_h whereas the probability to remain in the low state is p_l.

Fundamentalist still estimates the misspecified model $x_t = \varphi_f + \pi_f y_t$

5.4 Results

I simulate an economy where bot misspecification are present. I calibrate the markov chain to have a structural breaks every 100 periods in average. α_l and α_h are three percent deviation from the average value of α. I set σ_v atv the same level than σ_u but introduces a small persistance coefficient with $\rho_v = 0.3$. I display the convergence with respect to $(\nu, \lambda$ in figure (8). The two misspecification significantly enhances the dominance of the follower equilibrium.

5.5 Why do fundamentalists misspecify their model

It seems implausible to assume that fundamentalists will not detect the misspecification. However, I have two reasons to keep assuming they estimate the misspecified model.

First, I consider small deviations from the original model. For example, α_l and α_h are three percent deviation from the average value of α and the autoregressive coefficient ρ_v is only 0.3.

Second, even if they detect a misspecification, they could have serious trouble to identify and estimate the true model. In case of structural breaks, they should estimate no less than five parameters $\pi_f, \alpha_l, \alpha_h, p_h, p_l$. The number is the same if there is an unobserved variable $(\sigma_v, \sigma_u, \rho_v, \varphi_f, \pi_f$. If both misspecification are present, They have 8 parameters to estimate. In every case, they still observe two variables.

6 Applications

6.1 Cobweb model

The first example is a cobweb model. There are two reasons for that choice. First, this is the standard example in the expectation literature and especially the learning one. Second, the model is very close to the framework I describe in the previous section.

Cobweb models were introduced by Ezekiel (1938). They outline a partial equilibrium problem. Demand is determined by current prices but producers have to decide supply in advance. Thus, their production depends on their expectation of the equilibrium price. Our presentation heavily borrows from Evans and Honkapoja (2004).

The demand for the good in period t is

$$D_t = DP_t^{-a}W_t^{\phi}(1+\epsilon_t) \tag{22}$$

P_t is the price. W_t is an exogenous variable and ϵ_t is a white noise. W follows an autoregressive process. a and ϕ are parameters.

$$W_t = W_{t-1}^{\theta_w}(1+\epsilon_t^w)$$

ϵ_t and ϵ_t^w are white noise. Their standard deviation are σ and σ_w

Supply of an individual producer S_u depends on the price expected by this individual producer.

$$S_{i,t} = SP_{i,t}^{b,e} \tag{23}$$

b is a parameter.

Assume that every producer have the same expectation.

The market clearing condition gives

$$P_t = \frac{D}{S}^{\frac{1}{a}} W_t^{\frac{\phi}{a}} P_{e,t}^{\frac{-b}{a}}(1+\epsilon_t) \tag{24}$$

I rewrite the equation in logs.

$$p_t = log(cte) + \frac{\phi}{a}w_t - \frac{b}{a}p_{e,t} + \epsilon_t \tag{25}$$

The equation is similar to equation (1), p being equivalent to x and w to y.

Rational expectation solution I compute the expected value of the log of the equilibrium price

$$E(p_t) = \frac{a}{a+b}log(cte) + \frac{\phi}{a+b}w_t$$

$log(P_t)$ is gaussian thus the expected value of P_t is

$$E(P_t) = cte^{\frac{a}{a+b}}W_t^{\frac{\phi}{a+b}}e^{\sigma^2/2}$$

I deduce the rational expectation price

$$P_t = cte^{\frac{a}{a+b}}W_t^{\frac{\phi}{a+b}}e^{\frac{-b\sigma^2}{2a}}$$

Algorithm Algorithm is very close to our baseline case.

The only slight difference is that agents form a forecast for the log price. I propose the following variant for the algorithm.

1. Using the model they have chosen, agents compute the forecast for $p_t = logP_t$
2. They compute the expectation of P_t using log and multiplying by $e^{\sigma^2/2}$. σ is supposed to be known by agents.
3. With these forecasts, their individual supply is computed using equation (19)
4. The equilibrium price is computed
5. The log of the equilibrium price is computed
6. Agents compare the log of the equilibrium price.
7. If their model underperforms, they switch to the other model with probability μ
8. Once they have chosen their new model, they estimate it using the hostory of log equilibrium prices and value for exogenous variables

6.2 Asset price model

The second model I consider is an asset price model. The demand for an asset depends on its expected return. This is not a true asset price model in the spirit of Lucas(1978). However, the equilibrium condition is similar to a valuation equation.

Environment Consider an asset generating a dividend d_t

The dividend evolves according the process

$$d_{t+1} = \alpha + \lambda_1 d_t + \lambda_2 s_{t+1} + \epsilon_{t+1} \tag{26}$$

d_t is the dividend of the previous period. s_{t+1} is an exogenous variable which affects the dividend in $t+1$ along past dividends. s also follows an autoregressive process and is observable

$$s_{t+1} = \rho_s s_t + \epsilon^s_{t+1}$$

The supply of assets is inelastic and equal to 1. Demand by individual investors are linear function of the expected return

$$D^i_t = \frac{P^{e,i}_{t+1} + d^{e,i}_{t+1}}{P_t(1+r)} \tag{27}$$

$P^{e,i}_{t+1}$ and $D^{e,i}_{t+1}$ are individual expectations. r is the exogenous constant interest rate.

The mass of investors is equal to 1. The market clearing condition gives the usual equation for asset valuation in a deterministic framework

$$P_t = \frac{P^e_{t+1} + d^e_{t+1}}{1+r} \tag{28}$$

P^e and d^e are the sum of individual expectations.

An investor wants to forecast d_{t+1} and P_{t+1} in period t. He does not know the parameters of the different autoregresive process.

In period t, Investors form their expectation before going to the market. They observe P_{t-1}, d_t and s_t. Using this information, they make a forecast for P_{t+1} and d_{t+1}. Next, they go to the market. Once on the market, they adjust their demand with the asset price P_t but the knowledge of $P_t$do not modify their forecasts for P_{t+1}

Rational expectations solution The RE solution implies that the control variable (prices) is a function of state variables. I use the method of undetermined coefficients to compute it.

$$P_t = \phi_0 \phi_1 d_t + \phi_2 s_t$$

with

$$\begin{aligned}
\phi_0 &= \frac{\alpha}{1+r-\alpha} \\
\phi_1 &= \frac{\lambda_1}{1+r-\lambda_1} \\
\phi_2 &= \frac{\lambda_2 \rho_s}{1+r-\lambda_2 \rho_s}
\end{aligned}$$

Heterogenous expectations The model differ more widely than the cobweb model from our baseline framework. Agents have to forecasts both future dividends and future prices and they make forecasts for the period $t+1$ and not the period t. The adapatation of the algorithm is not so complicated

dividends I assume both followers and fundamentalists learn the "right" model for dividends. They estimate

$$d^e_{t+1} = \theta_0 + \theta_1 d_t + \theta_2 s_t + \varepsilon^d_t \tag{29}$$

Information about s_t and d_t is available and there is no feedback effect from dividend expectation to dividend themselves. Thus, there are no reason for an agent not to use all information to forecast future dividends.

Prices expectations Followers forecasts future prices by using past ones. They estimate

$$P_{t+1} = \psi P_t + \varepsilon_t^{p,o} \tag{30}$$

And their forecasts for P_{t+1} is

$$P_{t+1}^o = \psi^2 P_{t-1} \tag{31}$$

Fundamentalists estimate the model

$$P_{t+1} = \pi_1 d_t + \pi_2 s_t + \varepsilon_t^{p,f} \tag{32}$$

Their forecast for $t+1$ is

$$P_{t+1}^f = \pi_1 d_t + \pi_2 s_t \tag{33}$$

Algorithm I propose the following variant of the algorithm

1. Using the model they have chosen, their past etimates of parameter values and the value of d_t, s_t, p_{t-1} agents compute the forecast for D_{t+1} and P_{t+1}
2. The equilibrium price is computed
3. Equilibrium price in t is computed
4. actual price P_t is compare with their past forecasts P_t^e which have been made in $t-1$
5. If their model underperforms, they switch to the other model with probability μ
6. Once they have chosen their new model, they estimate it using the history of equilibrium prices and value for exogenous variables

A note on literature

6.3 Macroeconomic model

The last model I simulate is a very simple keynesian model.

Households, expectation and consumption Consumers maximize the utility function

$$
\begin{aligned}
& U(C_t^i, \omega_{t+1}^i) \\
w.r.t \;\; & C_t^i + Q_t(a_{t+1}^i)\omega_{t+1}^i = \omega_t^i + a_t^i
\end{aligned}
$$

Where a is the financial wealth a and ω is the labor wealth.

$$\omega_t^i = w_t E_t + \sum_{k=t+1}^{+\infty} \left(\prod_{j=t+1}^{k} Q_{e,j} \right) (w^{e,k} E_{e,k})$$

w is the wage and E is the employment level. Both are given for households. Q is the inverse of the interest factor When a variable has a small subscript e, it means that is a forecast and not a certain variable. The

$$Q_t = \frac{1}{R_{t+1}} \tag{34}$$

This way to write the consumer problem is unconventionnal but actually carry little difference with the standard problem. The advantage is that it allows an explicit derivation of demand function for a large class of utility function. These demand function for leisure and consumption can be expressed with respect to expected wages and expected interest rate. Both expected wages and expected interest rate are function to future income. The relation between income and real interest rate on one side and between income and wages on the other is supposed to be common knowledge among agents.

The utility function I consider is

$$U(c_t, \omega_{t+1}) = c_t^{\theta}(a_{t+1} + \omega_{t+1})^{1-\theta} \tag{35}$$

The cobb douglass utility function allows a straightforward derivation of consumption function

$$C_t^i = \theta(\omega_t^i + a_t^i) \tag{36}$$

Because of the linearity, aggregation is straightforward

$$C_t = \int_0^1 (\theta a_t^i + \theta \omega_t^i) = a_t + \theta w_t E_t + Q_t \theta \int_0^1 \omega_{t+1}^{e,i}$$

To make their consumption choice, consumer have to forecast the whole path of employment and interest rate.

The real interest rate is pegged according to the central bank rule

$$R_{t+1} = R^{nat} \left(\frac{Y_t}{Y_t^n} \right)^{\phi} e^{m_t} \tag{37}$$

m_t is a monetary policy shock. It follows an autoregressive process. R^{nat} is the real interest rate which allows full employment at steady state.

$$m_t = \rho_m m_{t-1} + \epsilon_t^m$$

The monetary policy rule is common knowledge among agents and they know ϕ

Productive sector Output Y_t is proportionnal to the employment level E_t.

$$Y_t = E_t \tag{38}$$

Desired level of employment by households is equal to one. On the short run, they have to serve labor demand by firms.

the natural output is equal to an employment equal to unity

$$Y_t^n = 1 \tag{39}$$

The interest rate is thus linked to employment

$$R_{t+1} = E_t^{\phi} e^{m_t} \tag{40}$$

Wages are equal to the marginal product of labor.

$$w_t = 1 \tag{41}$$

Return to scale are constant, so there are no profits.

Aggregate demand is equal to consumption.

$$Y_t = C_t \tag{42}$$

There are no goverment or corporate bond. The aggregate supply of financial asset is equal to zero and the market clearing condition is

$$a_{t+1} = 0 \tag{43}$$

Rational Expectation

Steady state At steady state,

$$Y = 1$$
$$\omega = \frac{1}{1-Q} = \frac{R}{R-1}$$

The natural interest rate is given by the equality between aggregate demand and aggregate supply

$$1 = \theta \frac{R}{R-1} \Rightarrow R^{nat} = \frac{1}{1-\theta}$$

deviation from steady state I write the model in term of deviation from steady state. It can be summarized by the three followin=g equations.

$$\tilde{e}_t = \theta \tilde{\omega}_t \tag{44a}$$
$$\tilde{r}_{t+1} = \phi \tilde{e}_t + m_t \tag{44b}$$
$$\tilde{\omega}_t = \theta \tilde{e}_t - (1-\theta)\tilde{r}_{t+1} + (1-\theta)\tilde{\omega}_{t+1} \tag{44c}$$

The deviation from steady state can be expressed with respect to m_t

$$\tilde{e}_t = \frac{\theta(\theta-1)}{1-\theta^2+(1-\theta)\theta\phi-(1-\theta)\rho_m} m_t$$
$$\tilde{\omega}_t = \frac{\theta-1}{1-\theta^2+(1-\theta)\theta\phi-(1-\theta)\rho_m} m_t$$
$$\tilde{r}_{t+1} = \phi\theta \frac{\theta-1}{1-\theta^2+(1-\theta)\theta\phi-(1-\theta)\rho_m} m_t + m_t$$

Heterogenous expectation Agents are supposed to know steady state values in this model. However, they are not able to connect these steady state values for endogenous variables with deep parameters. They also know the monetary policy process (the link between real interest rate and employment) and the process followed by monetary policy shock (they know ρ_m

Moreover, in t agents only observe their own employment at the period t and the aggregate employment of the previous period. They do not aggregate employment at period t.

Both type of agents have to forecast the whole path of employment.

Employment forecasts Fundamentalists estimate the equation

$$\tilde{e}_k = \pi_f m_k + \varepsilon_k^f \tag{45}$$

In period t, with the estimate $\pi_{m,t}$, they form a forecast for $\tilde{e}_{t+1}$

$$\tilde{e}_{t+1}^f = \pi_{f,t}\rho_m m_t \tag{46}$$

Followers estimate the equation

$$\tilde{e}_k = \pi_o \tilde{e}_{k-1} + \varepsilon_k^o \tag{47}$$

In period t, with the estimate $\pi_{m,t}$, they form a forecast for $\tilde{e}_{t+1}$

$$\tilde{e}_{t+1}^o = \pi_{o,t}^2 \tilde{e}_{t+1} \tag{48}$$

From employment forecasts to wealth estimation The main departure from the two previous models is that there is now an intermediate step between forecasts and policy rules.

Given their forecasts for employment, agents have to compute their labor wealth in $t+1$.

Fundamentalists compute their labor wealth according to

$$\tilde{\omega}_{t+1}^f = \frac{(1-q)\pi_{f,t} - q\phi\pi_{f,t}\rho_m - q\rho_m}{1-q\rho_m^2} m_t \tag{49}$$

Followers use the equation

$$\tilde{\omega}_{t+1}^o = \frac{(1-q-q\phi)\pi_{o,t}^2}{1-q\pi_{o,t}^3}\tilde{e}_{t-1} - \frac{q\rho_m}{1-q\rho_m^2} m_t \tag{50}$$

Both knows steady state values ω^* and other steady state values. Thus, they are able to compute their expected wealth, respectively $\omega_{f,t+1}$ and $\omega_{o,t+1}$

The equilibrium equation is

$$E_t = \theta E_t + \frac{\theta}{1-\theta} E_t^\phi e^{m_t} \gamma_t \omega_{t+1}^o + \frac{\theta}{1-\theta} E_t^\phi e^{m_t}(1-\gamma_t)\omega_{t+1}^f \tag{51}$$

7 Literature

Rational expectations and learning literature Rational expectations were introduced by Muth (1961). The learning literature was developed to adress the issue of limited knowledge of parameter values. A classical exposition can be found in Evans and Honkapoja (2001). Convergence theorem are due to Bray and Savin (1986), a result refined by Marcet and Sargent (1989)

Rational expectations assessment: sunspots Rational expectations were challenged by the sunspot literature initiated by cass and Shell (1977) and refined by Azariadis and Guesnerie (1982). These two papers have shown that, in some class of models, exogenous variables completely unrelated to endogenous variable may affect them simply because agents believe they do. Our idea is quite close. Lagged endogenous variable does not affect directly current ones but may through beliefs. The difference is that lagged endogenous variable are correlated to current ones through the persistence of fundamental exogenous variables. Intuitively, they may play a role in a larger class of models whereas pure sunspots needs strict conditions to emerge (see Guesnerie 2001 for a review).

Evolutionnary theory and economics: Nelson and Winter, Saint Paul The evolutionnary viewpoint has a long history in economics. Some intuitions may be found in Schumpeter (1926) and in the austrian school. Friedman (1953) has defended the rationality assumption by suggesting that "rational" agents will eliminate "irrationnal" ones in markets. The volutionnary viewpoint was formalized in a more rigorous way by Nelson and Winter (1982) and more recently by Saint Paul (2015).

Agent based expectations: Brock and hommes Our paper is more directly related to three approaches. The first one is the Consistent Expectation Equilibrium (CEE) literature developed by Brock and Hommes since their seminal paper (1997) and refined in a recent textbook by Hommes (2012). Consistent Equilibrium Expectation departs from rational expectation by imposing much weaker condition for expectations. Expectations should simply be consistent with observed autocorrelations. The link with the behavior of our followers is obvious. In the Brock and Hommes original paper, agents switch between rational and naive expectations according to a performance/cost comparison.

There are however several difference between my paper and this branch of the literature. I am interested by the convergence toward one type of expectations, either rational expectations or more adpatative ones. Brock and Hommes (1997, 1998) and Hommes (2012) are more interested by the cyclical dynamic, or even the chaotic one, induced by the coexistence of naive and rational expectations.

A second important difference is the learning behavior. Learning is very simple in most of the CEE literature. In Hommes (2012), the equivalent of our followers are endowed with *given* forecasting rules and do not learn parameters of the forecasting rules. Fundamentalists know the true model of the economy and have not to learn it. By contrast, Our approach heavily borrows from Evans and Honkapoja learning.

A third difference is the evolutionnary criteria. In Brock and Hommes, agents use a discrete choice model to choose between the different forecasting rules. I choose a more intuitive criteria which also allows for convergence toward one rule more easily.

Agent based modelling Close to the CEE literature is agent based modelling. Agent based models takes the opposite approach to standard theory. Instead of deriving the optimal behavior from a well defined maximization problem, they impose given behavioral rules to agents. A good example of agent based model for asset market can be found in Lebaron (2005). The drawback of this approach is the high number of possible behavioral rules and the large degree of freedom it gives to the modeller. Our contribution is closer to the standard theory. I look at behavioral rules which can outperform "rational" ones in environment with limited knowledge.

Adam-Marcet asset prices Another paper closely related to mine is a recent paper by Adam and Marcet (2011). In this paper, they compare two learning strategies in a Lucas asset market. In the first one, agents learn the relation between price and current dividends. In the second one, agents learn the relation between current and past prices. they show that the second learning strategy offers a simple explanation to many asset pricing puzzle. There are two differences with my paper. First, in their model both strategies converges to the rational expectation solution. This is because dividends follows a very simple process in which past dividends are a sufficient statistics to forecast future ones. Because with rational expectations past prices are function of past dividends, past prices are also a sufficient statistics for future asset price. If dividend equation are more complex, past prices are no more a sufficient statistics to forecast future ones. Agents miss available information. Thus, we are not convinced that the second learning strategy proposed by Adam and Marcet is compatible with rational expectations in a more general setup.

A second difference is the role of evolution in our model. Adam and marcet compares the ability of two learning models to explain stylized facts. We look at the selection process between different learning strategies.

8 Conclusion

References

[1] **Adam, K. and A. Marcet.** 2011. "Internal rationality, imperfect market knowledge and asset prices." *Journal of Economic Theory* 146,1224-1252

[2] **Bray, M., and N. Savin.** 1986. "Rational expectations equilibria, learning, and model specification." *Econometrica* 54,1129-1160

[3] **Brock, W.A. and C.H. Hommes.** 1997. "A rational route to randomness." *Econometrica* 65,1059-1095

[4] **Bullard, J. and J. Suda.** 2011. "The stability of Macroeconomic systems with Bayesian learners" *Banque de France working papers*

[5] **Cass, K. and A. Shell.** 1983. "Do sunspots matter" *Journal of Political Economy* 91,193-227

[6] **Evans, G., and S. Honkapoja.** 2001. *Learning and expectations in Macroeconomics* Princeton University Press

[7] **Friedman, M.** 1957. *A Theory of the Consumption Function.* Princeton University Press.

[8] **Grandmont, M.** 1998. "Expectations Formation and Stability of Large Socioeconomic Systems." *Econometrica* 66,741-781

[9] **Guesnerie, R.** 2001. *Assessing Rational Expectations: Sunspot Multiplicity and Economic Fluctuations.* MIT Press

[10] **Guesnerie, R.** 2005. *Assessing Rational Expectations 2-"eductive" stability in economics.* MIT Press

[11] **Guesnerie, R. and C. Azariadis** 1982. "Propheties creatrices et persistance des theories" *Revue Economique* 33,787-806

[12] **Guesnerie, R. and M.Woodford** 1991. "Stability of Cycles with Adaptive Learning Rules." in *Equilibrium Theory and Applications* ed. by W. Barnett, B. Cornet, C. d'Aspremont, J. Gabszewicz, and A. Mas-Colell. Cambridge University Press.

[13] **Hommes, C.** 2013. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems.* Cambridge University Press.

[14] **Hommes, C. and G.Sorger** 1998. "Consistent Expectations Equilibria." *Macroeconomic Dynamics* 2,287-321

[15] **Marcet, A., and J.P. Nicolini.** 2003. "Recurrent Hyperinflation and Learning." *American Economic Review* 93(5),1476-1498

[16] **Marcet, A., and T. Sargent.** 1989. "Convergence of least squares learning mechanisms in self referential linear stochastic models." *Journal of economic Theory* 48,337-368

[17] **McGough, B.** 2003. "Statistical learning with time varying parameters." *Macroeconomic Dynamics* 7,119-139

[18] **Muth, J.** 1961. "Rational expectations and the theory of price movements" *Econometrica* 29,315-335

[19] **Nelson, R. and S.G. Winter** 1982. *An Evolutionary Theory of Economic Change.* Harvard University Press.

[20] **Saint Paul, G.** 2014. "How Darwinian should an economy be?" PSE working papers.

[21] **Woodford, M.** 1990. "Learning to believe in sunspots". *Econometrica* 58, 277-307.

A Multiple equilibra

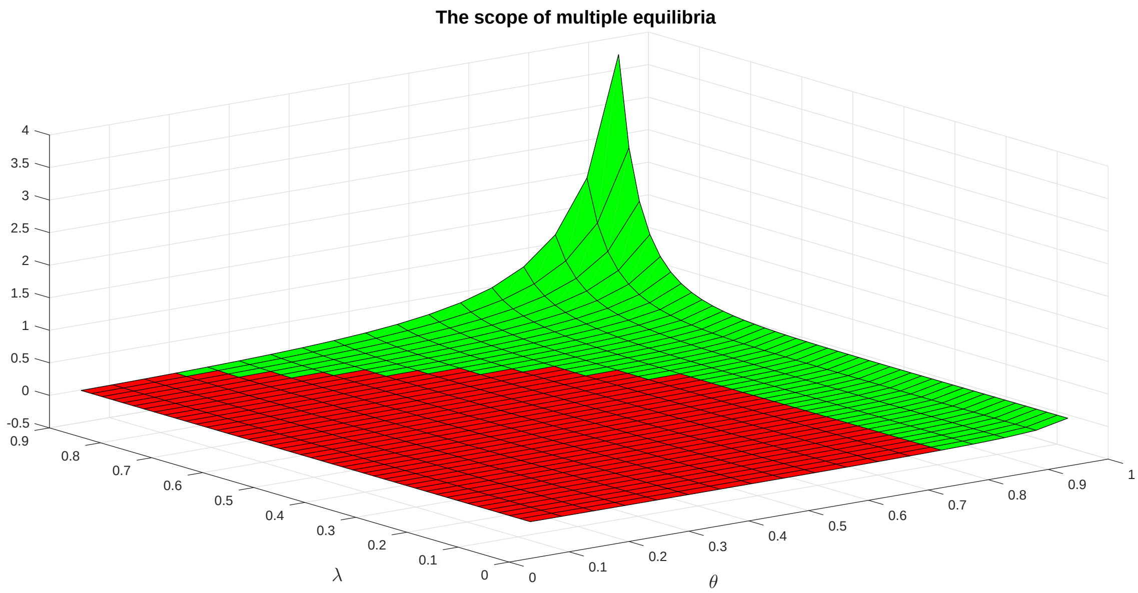

Figure 1: Convergence for several parametrization

B Calibration tables baseline model

Variable	Baseline	Fundamentalist	Follower
λ	0.65	0.25	0.65
ν	0.8	1	0.6
γ_0	0.5	0.5	0.5
θ	0.8	0.8	0.8
σ_y	0.01	0.01	0.01
σ_u	0.005	0.005	0.005
α	0.5	0.5	0.5
β	1	1	1

Table 1: calibration

C Results baseline model

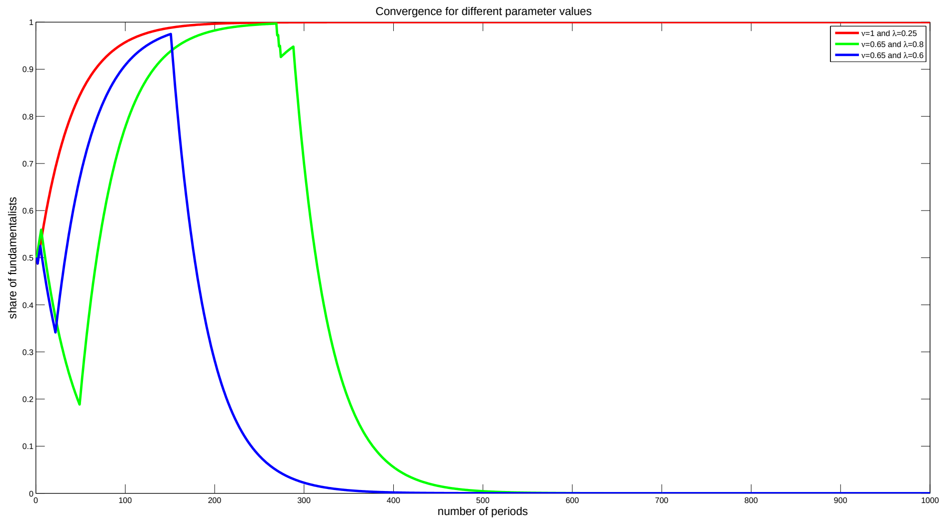

Figure 2: Convergence for several parametrization

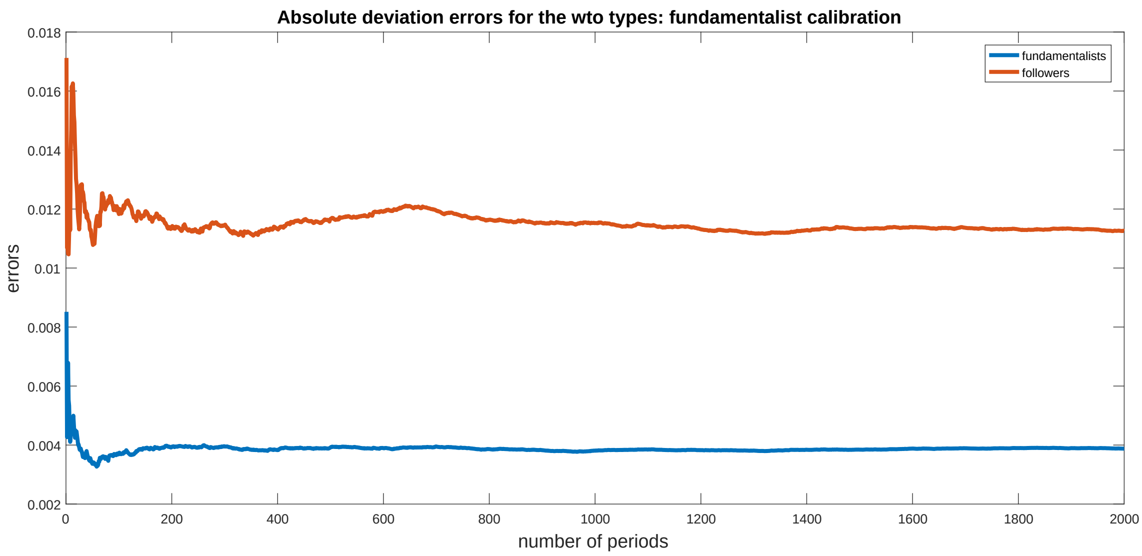

Figure 3: Errors for "fundamentalist" calibration

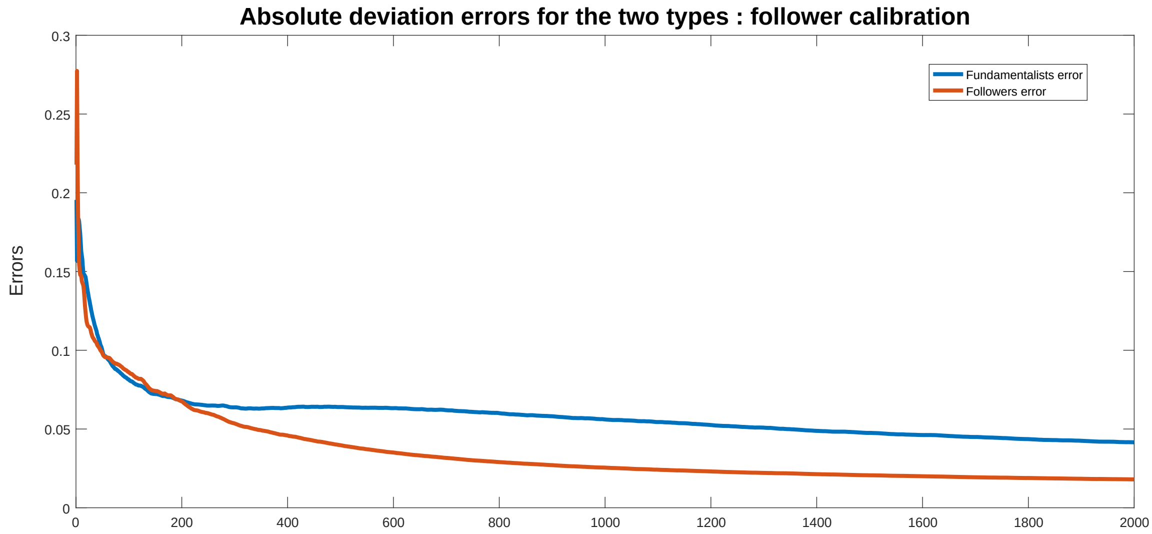

Figure 4: errors for followers calibration

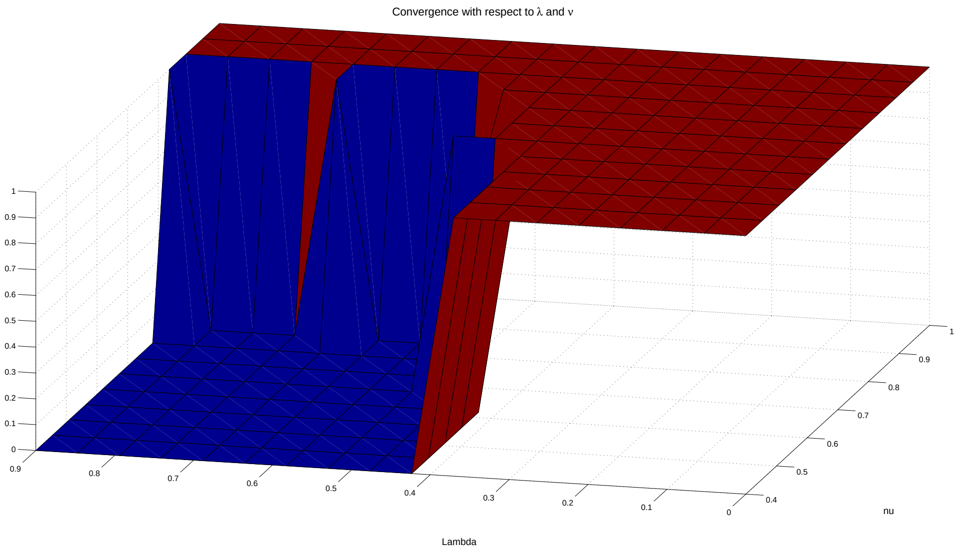

Figure 5: Sensitivity of the response to λ and ν

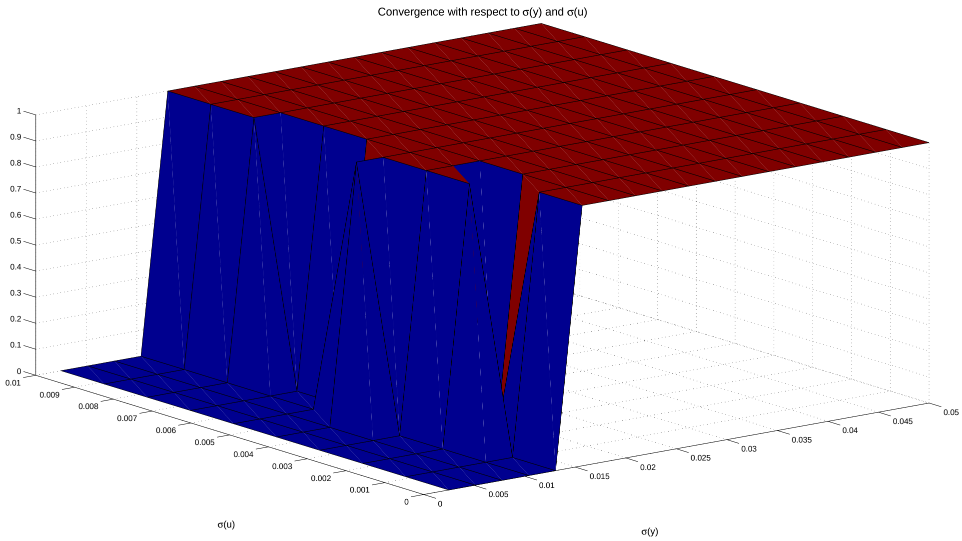

Figure 6: Sensitivity of the response to σ_y and σ_u

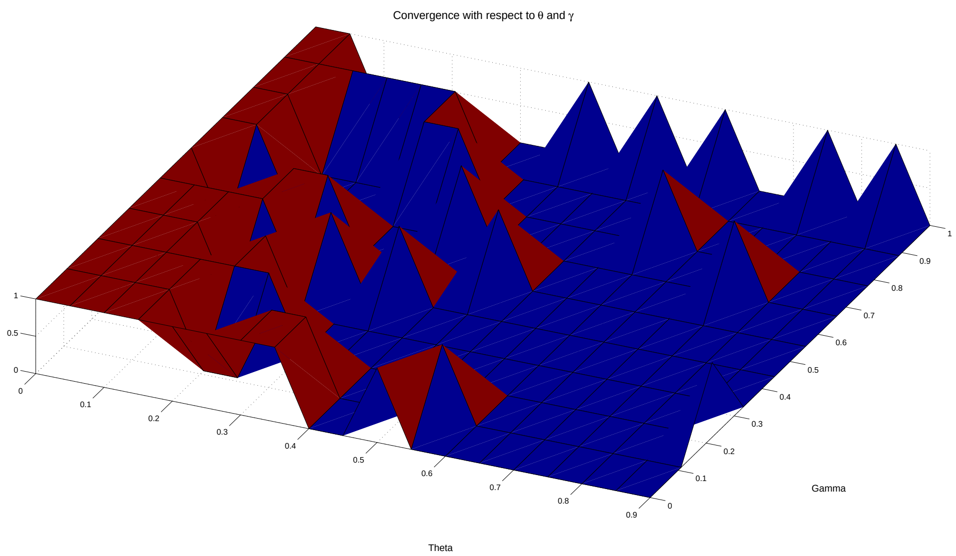

Figure 7: Sensitivity of the response to γ_0 and θ

D Results for structural breaks and unobervables

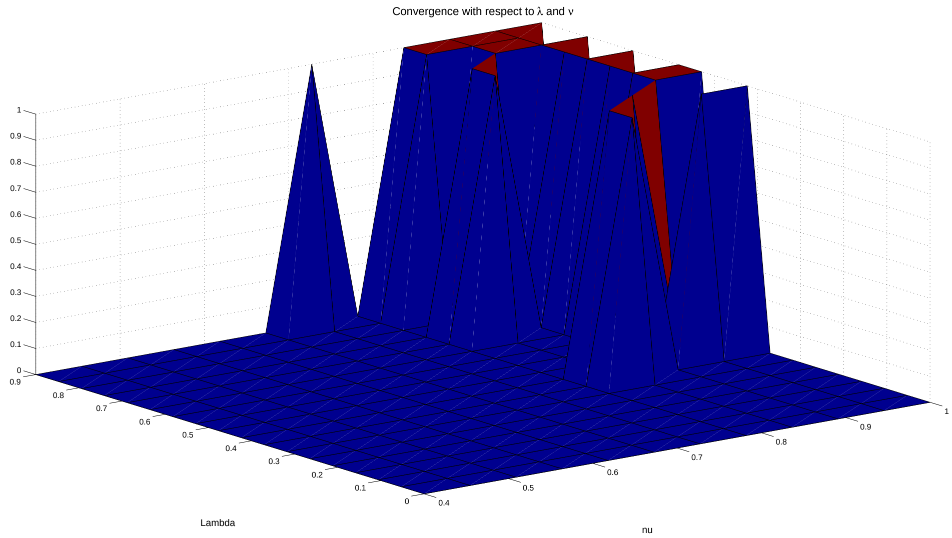

Figure 8: Sensitivity of the response to λ and ν

E Calibration tables

E.1 Asset prices

E.2 Keynesian model

F Results economic models

F.1 Asset prices

F.2 Keynesian model