Heterogeneous agents in the FX market: A matter of time horizon?*

Preliminary and incomplete, please do not quote or circulate

Sophie Béreau[†]

December 31, 2011

Abstract

This paper deals with a Bayesian extension of a behavioral finance framework 'à la' De Grauwe and Grimaldi (2006*a*) in which agents operating in the FX market differ in their forecasting time horizon for the exchange rate. More specifically, as well documented by recent surveys on trading behaviors, agents may target either the short, the medium or the long run, and thus potentially different equilibrium values for the exchange rate as shown by Bénassy-Quéré, Béreau and Mignon (2010). In the short run, if we believe in the world described by Meese and Rogoff (1983), this leads to a chartist rule, whereas in the long run, the PPP condition appears as a natural anchor. In between, i.e. in the medium run, we implement an APEER model using Bayesian tools, as an alternative to the FEER-BEER nexus. Our results show that the stabilizing impact of the intermediate rule depends on agents' good perception of the fundamentals.

JEL Classification: F31, C11

Keywords: Behavioral finance modeling, equilibrium exchange rate, Kalman filter, Bayesian analysis

1 Introduction

Numerous approaches aim to characterize exchange rate behavior. From an historical perspective, in the aftermath of the Bretton-Woods collapse in 1971, the first consistent modeling of exchange rate dynamics comes with the monetary models [1] which rely on the idea that exchange rates reflect the arbitrages made on both the financial and monetary markets. In practice however, those approaches fail at doing better in predicting short-run exchange rates than a simple random walk model, as the celebrated paper of Meese and Rogoff (1983) suggests. Thus, other research avenues have been explored.

*I would like to thank Jean-Yves Gnabo, Paul De Grauwe, Giulia Piccillo, Vincent Bouvatier, Guy Laroque, Jean-Michel Grandmont, Etienne Lehmann, Frédéric Bec and seminar participants at CREST, Thema - Université de Cergy-Pontoise, ADRES, University of Kent and Université Paris Ouest, for their very helpful comments and insights on this work in progress.

†Université catholique de Louvain - Louvain School of Management, Office a.204, Box L2.01.02, Place des Doyens, 1, 1348 Louvain-la-Neuve, Belgique. E-mail: sophie.bereau@uclouvain.be

1. We can quote the contributions of Frenkel (1976), Bilson (1978), Dornbusch (1976), Frankel (1979), or Hooper and Morton (1982) for the most famous of them.

On the one hand, following Obstfeld and Rogoff's seminal contribution [2] on the "Exchange Rate Dynamics Redux" numerous theoretical contributions have tried to propose microfounded views of exchange rate dynamics, leading to the so-called "New Open Economy Macroeconomics" area. As De Grauwe and Grimaldi (2006*a*) notice, those developments have been pursued without any empirical relevance considerations, leading to various "puzzles", as defined by the same pioneering authors. [3] Alternative views have then emerged to explain the occurrence of such puzzles. Among them we can quote the "microstructure" literature [4] that gives some insights to better understand exchange rate behavior in the (very) short run and "behavioral finance" models [5] which, by relaxing the assumption of agents' full rationality, engender complex dynamics that are much closer to what is currently observed in FX markets, both in the short and the long run. More specifically, as Hommes (2008) explains, in such a framework, financial markets are viewed as complex, adaptive systems consisting of many boundedly rational agents interacting through simple heterogeneous investment strategies, constantly adapting their behavior in response to new information, strategy performance and through social interactions.

On the other hand, many empirical contributions that aim at assessing equilibrium values for the exchange rate, from the PPP anchor to more advanced concepts, have developed largely aside from those theoretical considerations. Among those concepts we can distinguish two main approaches that are the FEER and the BEER models. The "FEER" refers to the "Fundamental Equilibrium Exchange Rate" and has been primary developed by Williamson (1985). It relies on the idea that the target exchange rate in the long run is the one that allows the realization of both internal and external *equilibria* in the economy. Turning to the "BEER", namely the "Behavioral Equilibrium Exchange Rate" model, it has been proposed by Clark and MacDonald (1998). More specifically, this approach consists in estimating a cointegration relationship between the exchange rate and a set of fundamentals. Equilibrium values are then derived from the estimated coefficients of the long-run equation and compared to the current ones, which allows to disentangle currencies' over- and undervaluation phases. In Bénassy-Quéré et al. (2010), we have shown that rather than being opposed, those different views of equilibrium exchange rates could be considered as *equilibria* of the same partial equilibrium model at different time horizons. This finding stems from the observation that exchange rates' macro determinants such as productivity differentials, stocks and prices do not adjust concomitantly.

In this paper, we try to reconcile those two strands of the literature by introducing the empirical expertise developed in Bénassy-Quéré et al. (2010) and Bénassy-Quéré, Béreau and Mignon (2009) into a behavioral finance model. As mentioned before, equilibrium exchange rate values may differ depending on the time horizon agents deal with. In a series of papers, Gehrig and Menkhoff (2003, 2004, 2005) show that rather than the usual opposition between "fundamentalist" and "chartist" agents as defined by Frankel and Froot (1990), one can distinguish "FX dealers", who act on a daily basis from "fund managers", who are more sensitive to longer time horizons. This opposition clearly paves the way to the investigation of exchange rate dynamics resulting from the opposition of different agents' trading strategies in terms of time horizon.

2. See Obstfeld and Rogoff (1995).
3. See Obstfeld and Rogoff (2000).
4. See Lyons (2001) for an overview.
5. See De Grauwe and Grimaldi (2006*a*) for an extensive development of this framework on the exchange rate issue and more recently Piccillo (2009) or Rovira Kaltwasser (2010) for significant extensions.

- In the short run: It may be consistent for investors to stick to a "chartist" forecasting rule, since this type of behavior appears rather successful at a time horizon for which the relevance of fundamental determinants is questioned.
- At longer time horizons: Fundamentals may play a key role, which leads investors to potentially rely on more sophisticated "fundamentalist" forecasting rules.

Relying on a behavioral finance model "à la" De Grauwe and Grimaldi (2006*a*), we then assume three different forecasting rules depending on the time horizon targeted by the agents operating on the FX market. [6] As previously mentioned, the short run approach may refer to the "chartist" rule. According to Meese and Rogoff (1983), the best predictor in the short-run is the random walk model, implying that neither fundamentals nor the past of the series can help to predict the exchange rate. In practice however, this vision seems a bit restrictive. It has been proved both in survey analyses (Taylor and Allen, 1992; Cheung, Chinn and Marsh, 2004) and theoretical contributions (Bask, 2007), that technical trading is a common practice especially in the short run when the impact of fundamentals is questioned. At longer time horizons, we assume two different frameworks for the assessment of the "equilibrium exchange rate" that will be used in the medium and long run forecasting rules respectively. Whereas in the long run, the PPP value appears as a natural anchor, as previously mentioned, various models exist to assess "equilibrium" values in the medium run for the real exchange rate. [7] We propose an "atheoretical" specification in the sense of MacDonald (2000) as an alternative to the FEER-BEER nexus, which consists here in the filtering of the exchange rate series through a Kalman filter as proposed by Engel and Kim (1999). We then use Bayesian tools to derive our estimations of the medium-run equilibrium exchange rate value. We do believe that this framework is perfectly accurate to assess how agents' ability to correctly perceive *ex-ante* the fundamentals impacts exchange rate dynamics. [8]

The remainder of the paper is the following. Section 2 specifies the full behavioral framework we rely on to model the exchange rate. We first describe the standard approach of De Grauwe and Grimaldi (2006*a*) in which we include an intermediate forecasting rule in sub-section 2.1. We then fully specify the state-space representation and derive the Kalman filter for the medium-run model in sub-section 2.2. In Section 3 we draw and comment representative simulations. Finally, Section 4 proposes some preliminary analysis and Section 5 concludes.

6. The analysis of more than two forecasting rules in such a behavioral model has already been explored by De Grauwe and Rovira Kaltwasser (2007) among others, but here what drives the investor's choice among various forecasting rules is not his mood (i.e. optimistic vs. pessimistic rule in presence of uncertainty on the fundamental value) but his willingness to target the exchange rate at different time horizons as documented in Gehrig and Menkhoff (2003, 2004, 2005).

7. See Driver and Westaway (2004) for a review.

8. Note that the impact of agents' uncertainty about fundamentals has already been explored in De Grauwe and Rovira Kaltwasser (2007) or Rovira Kaltwasser (2010) using advanced techniques from Physics. Here is an attempt to introduce such a problematic in a Bayesian context.

2 The exchange rate model

2.1 A behavioral finance approach

In this section, we present a simple nonlinear model for the exchange rate, in line with the behavioral finance framework described in Chapter 2 of De Grauwe and Grimaldi (2006*a*)[9] in which we include an intermediate forecasting rule that will be fully specified in sub-section 2.2.

In this model, the exchange rate dynamics stems from the confrontation of agents' potentially heterogeneous beliefs on its future value. More specifically, agents operating on the FX market make a forecast of the future exchange rate by means of a "rule of thumb" forecasting rule at the beginning of the period. After each period, they assess the relative gain/loss associated to their *ex-ante* choice of selecting a particular rule using a fitness criterion, i.e. by comparing the *ex-post* risk-adjusted profit of their rule with the average *ex-post* risk-adjusted profits of alternative rules. Agents may then consider switching to a more profitable rule in a rational manner. Let us first describe what kind of agents we deal with and how the confrontation of their heterogeneous predictions leads to our specification of the future exchange rate.

As previously mentioned, we assume a constant number of agents potentially of three types according to their forecasting time horizon, which can be either the short, the medium or the long run. To these various time horizons, correspond different equilibrium values for the targeted exchange rate. If we believe in the world described by Meese and Rogoff (1983), the best short-run predictor for the exchange rate is the observed, current value, implying that neither fundamentals nor the past of the series are of great help to forecast the future movements of the exchange rates. However, it has been well documented both in the theoretical and empirical literatures that at that time horizon, FX dealers relied on some form of technical analysis for their forecasts due to their good performance.[10] Agents who target the short run will then have in mind that to forecast the series they may rely on its immediate past, i.e. the positions of the last few hours or days, which is exactly what "chartists" do. At longer time horizon, fundamentals may play a key role.[11] We then assume two distinct anchors for the exchange rate, which are the PPP, denoted by $\overline{s}_{l,t}$,[12] in the long run and the filtered value of a structural model, denoted by $\overline{s}_{m,t}$, that will be fully specified in sub-section 2.2, in the medium run. In addition, we assume that the PPP exchange rate follows a random walk without a drift,[13] which leads to:

$$\overline{s}_{l,t} = \overline{s}_{l,t-1} + \varepsilon_t,\, \varepsilon_t \sim i.i.d\, \mathcal{N}(0, \sigma^2_\varepsilon) \tag{1}$$

Short-run agents extrapolate past movements of the exchange rate into the future, they are then assumed to follow as chartist agents in De Grauwe and Grimaldi (2006*a*), a positive feedback rule such as:

$$\mathbf{E}_{s,t}(\Delta s_{t+1}) = \beta(\Delta s_t),\, \beta > 0 \tag{2}$$

9. Very similar results for the exchange rate dynamics may be obtained in a less simple framework, i.e. where agents choose first their optimal international portfolio relying on a standard mean-variance criterion, as presented in Chapter 3 in De Grauwe and Grimaldi (2006*a*), (pp. 49-69) and recalled in Appendix B.

10. See Neely (1997) for a short note on that practice.

11. See Mark (1995); Mark and Sul (2001) and more recently Christopoulos and León-Ledesma (2009) among others.

12. All the exchange rate series considered in this paper are expressed in logarithms.

13. This hypothesis is consistent with PPP holding in the long run. In the empirical literature, it is assumed that if the absolute PPP holds in the long run, then logs of price differences and the nominal exchange rate should cointegrate, which implies $(s_t, \overline{s}_{l,t}) \sim$ CI(1,1), see Froot and Rogoff (1995) for more details.

with $\mathbf{E}_{s,t}$ the forecast made by short-run agents using one-step past information and s_t the observed exchange rate. [14]

Regarding now agents that forecast longer time horizon and let us define $m_{m,t}$ and $m_{l,t}$ as the medium and long-run currency misalignments as follows:

$$m_{k,t} = s_t - \overline{s}_{k,t}, \forall k \in \{m, l\} \tag{3}$$

We get the following result. Observing the current value s_t, both medium and long-run agents are able to measure its misalignment compared to their respective equilibrium anchor $\overline{s}_{k,t}$. If for example, s_t is above $\overline{s}_{k,t}$, meaning that the associated misalignment is positive and that the currency is undervalued (if we assume the exchange rate to be expressed in terms of the domestic currency), then agents expect that the future exchange rate s_{t+1} will appreciate to reach $\overline{s}_{k,t}$ the next period, which leads to $\Delta s_{t+1} < 0$. They are thus following a negative feedback rule such as:

$$\mathbf{E}_{k,t}(\Delta s_{t+1}) = -\eta_k m_{k,t}, \text{ with } \eta_k > 0, \forall k \in \{M, L\} \tag{4}$$

At each period, agents may switch from one type to another depending on the *ex-post* performance of their forecasting rule. They use for that a fitness criterion inspired from Brock and Hommes (1997) assumed to be as follows:

$$\omega_{i,t} = \frac{\exp\left(\gamma \pi'_{i,t}\right)}{\sum\limits_{i \in \{s,m,l\}} \exp\left(\gamma \pi'_{i,t}\right)}, \forall i \in \{s, m, l\} \tag{5}$$

where $\omega_{i,t}$ is the share of the total population of agents relying on the rule of type i, $\gamma \geq 0$ the intensity with which agents revise their rules, $\pi'_{i,t}$ the risk-adjusted profit realized by means of the i's forecasting rule and s, m, l the short, medium and long-run respectively.

In this model, we assume the risk-adjusted profit to be defined as follows:

$$\pi'_{i,t} = \pi_{i,t} - \mu \sigma^2_{i,t} \tag{6}$$

with: $\pi_{i,t}$ representing the profit resulting from the use of rule i, $\sigma^2_{i,t}$ the risk incurring by agents when they make forecast that we assume as De Grauwe and Grimaldi (2006*a*) to be defined by the variance of the forecast error, and μ the coefficient of risk aversion.

We get:

$$\sigma^2_{i,t} = \left[\mathbf{E}_{i,t}(s_t) - s_t\right]^2, \forall i \in \{s, m, l\} \tag{7}$$

Following De Grauwe and Grimaldi (2006*a*), we define the profit $\pi_{i,t}$ of a certain rule i at time t, as the one-period earnings of investing one US$ in the foreign asset, which leads to:

$$\pi_{i,t} = \left[s_t\,(1 + r^*) - s_{t-1}(1 + r)\right] \cdot \text{sgn}\left[\mathbf{E}_{i,t-1}(s_t)(1 + r^*) - s_{t-1}(1 + r)\right] \tag{8}$$

14. De Grauwe and Grimaldi (2006*b*) consider a more complete structure of memory for the chartist agents', who are supposed to base their forecast on the entire past of the series. As this leads to very similar results we have stuck to the simplest approach that we believe to be more consistent with the rest of our assumptions, including that of limited access to information for medium-run agents, as described in Section 2.2.

with r, r^* designing the domestic and foreign interest rate respectively, and:

$$\text{sgn}(x) = \begin{cases} 1 \text{ for } x > 0 \\ 0 \text{ for } x = 0 \\ -1 \text{ for } x < 0 \end{cases}$$

Equation (8) stems from the following arbitrage: we assume that at time $(t-1)$ agents i may invest either 1 US$ in the foreign market or the same amount (in domestic currency) in the domestic one. At time t, the first operation leads to an amount of $(1+r^*)$ in dollars and thus to $s_t(1+r^*)$ in domestic currency, whereas the other strategy, which consists in investing s_{t-1} at time $(t-1)$ in the domestic market, leads to $s_{t-1}(1+r)$ at time t. Agents i then compare those two expected returns. They obviously choose the strategy that they expect to lead to an increase of their current wealth. Depending on their ability to correctly predict s_t in $(t-1)$, this amount may be a gain or a loss. More specifically, if agents have correctly predicted the future exchange rate s_t in $(t-1)$, then the signs of both factors are the same (either positive or negative) and lead to $\pi_{i,t} > 0$, which corresponds to a gain. In the opposite case, i.e. if agents have mispredicted s_t, then the signs of the two factors are opposite, which leads to $\pi_{i,t} < 0$ and corresponds to a loss.

Finally, assuming that the expected exchange rate variation is the weighted sum of expected exchange rate variations for all the agents operating in the FX market, we obtain:

$$\mathbf{E}_t(\Delta s_{t+1}) = \omega_{s,t}\beta(s_t - s_{t-1}) - \omega_{m,t}\eta_m(s_t - \overline{s}_{m,t}) - \omega_{l,t}\eta_l(s_t - \overline{s}_{l,t}) \tag{9}$$

The expression of the future exchange rate is then given by:

$$s_{t+1} = s_t + \omega_{s,t}\beta(s_t - s_{t-1}) - \omega_{m,t}\eta_m(s_t - \overline{s}_{m,t}) - \omega_{l,t}\eta_l(s_t - \overline{s}_{L,t}) + \widetilde{\varepsilon}_{t+1} \tag{10}$$

with: $\widetilde{\varepsilon}_{t+1} \sim \mathcal{N}(0, \sigma^2_{\widetilde{\varepsilon}})$ representing the news in time (t+1) that has not been incorporated in agents' forecasts.

Now we have detailed the structure of the behavioral model, we turn to the specification of the medium-run exchange rate model and its estimation procedure through a Bayesian analysis.

2.2 Dealing with equilibrium exchange rates at various time horizons

As showed in Bénassy-Quéré et al. (2010), equilibrium exchange rate values may differ according to the time horizon considered, calling for a more or less abrupt adjustment of the exchange rate. Thus, when agents have different time horizons in mind for their forecasts, they rely on potentially different equilibrium values for the exchange rate and the dynamics that are derived from their expectations may differ accordingly. As mentioned before, whereas in the short run, agents deal with a chartist-type rule since at that time horizon the past value is the best predictor for the future exchange rate, in the long run, we assume that the PPP theory holds. In between, a wide range of models exists. We propose an "atheoretical" assessment in the sense of MacDonald (2000) as an alternative to the FEER-BEER nexus. The rationale for using this approach is twofold. First, as argued by Driver and Westaway (2004), "Atheoretical Permanent Equilibrium Exchange Rate" or APEER models are assessed without any consideration of stock adjustment and as such, do correspond to a medium-run anchor. Second, as one of our aims here is to measure how medium-run agents' perception of fundamentals influence the exchange rate dynamics, we privileged a straightforward specification to get a clear message when interpreting Bayesian computations and leave for further developments the study of a more advanced modeling. More specifically, we follow Engel and Kim (1999) and derive our medium-run model by assuming a trend-cycle decomposition of the observed exchange rate series. We then use Bayesian techniques to implement the Kalman filter on the state-space representation. Filtered values of the permanent component provide medium-run *equilibria* for the exchange rate that will be finally introduced into the behavioral finance model described in Section 2.1.

2.2.1 An APEER specification

Relying on MacDonald (2000), the APEER model corresponds to the following decomposition of the real exchange rate:

$$q_t = q_t^P + q_t^C \tag{11}$$

with: q_t^P and q_t^C being the permanent and cyclical components of q_t, the real exchange rate defined as follows:

$$q_t \equiv s_t + p_t^* - p_t \tag{12}$$

with: p_t^* and p_t the log of foreign and domestic price indices respectively.

If we assume a stable level of inflation in both countries due for instance to Inflation Targeting (IT) policies, we get:

$$p_{t+1} = p_t + \alpha, \forall t \tag{13}$$

$$p_{t+1}^* = p_t^* + \alpha^*, \forall t \tag{14}$$

with α and α^* being small quantities. [15] If we assume that $\alpha = \alpha^* = 0$, then we get:

$$p_t^* - p_t = p_{t+1}^* - p_{t+1}, \forall t \Leftrightarrow p_t^* - p_t = c, \forall t \tag{15}$$

We can then assume $c = 0$ without lack of generality, which leads to:

$$q_t = s_t \tag{16}$$

As a result, targeting the real exchange rate q_t is equivalent to targeting the nominal one s_t in a world where agents assume inflation rates to be stable both in the home and foreign countries.

15. For instance, the figures corresponding to the percentage changes in CPI in the Eurozone and the US in April 2010 were of 0.2% and -0.1% respectively. Sources: ECB, Statistical Data Warehouse, and Bureau of Labor Statistics, United States Department of Labor. More generally, this assumption seems reasonable for OECD countries and emerging countries that exhibit lolw inflation targets.

In our framework, agents base their forecasts on the assessment of an equilibrium value for the exchange rate, assumed to be defined in the medium-run according to an APEER specification. This requires them to accurately disentangle permanent from cyclical factors in order to clearly identify s_t^P. In practice, numerous techniques have been developed in the literature to implement such a decomposition. A first approach is given by Beveridge and Nelson (1981) (BN hereafter) seminal contribution, that has lead to numerous empirical investigations, either in an univariate context as proposed by Huizinga (1987), or in a multivariate one, with studies from Cumby and Huizinga (1990) or Clarida and Gali (1994) among others. More recently, advanced concepts such as Gonzalo and Granger (1995) common factor decomposition that Clark and MacDonald (2004) have applied to a cointegrated-BEER equation have been explored. Here, we propose to implement the framework of Engel and Kim (1999) which consists in a simple univariate unobserved component model that we write in a state-space form and estimate through a Kalman filter. [16] The only difference with a BN decomposition is that we do not assume here that the shocks to the two components are correlated.

The model we assume can thus be written as follows:

$$\begin{aligned} s_t &= s_t^P + s_t^C && (17)\\ s_t^P &= s_{t-1}^P + \delta_{t-1} + \varepsilon_t,\, \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2) && (18)\\ \delta_t &= \delta_{t-1} + z_t,\, z_t \sim \mathcal{N}(0, \sigma_z^2) && (19)\\ s_t^C &= \phi_1 s_{t-1}^C + \phi_2 s_{t-2}^C + u_t,\, u_t \sim \mathcal{N}(0, \sigma_u^2) && (20) \end{aligned}$$

As described in Equation (17), we write the nominal exchange rate s_t as the sum of two unobservable components, a permanent one s_t^P, supposed to be the equilibrium value for the exchange rate in the medium run, and a cyclical one s_t^C, corresponding to the misalignment, i.e. the gap between the observed current value and the equilibrium one. This expression is then equivalent to Equation (3) for $k = M$, with $s_t^P = \overline{s}_{k,t}$ and $s_t^C = m_{k,t}$.

While s_t^P is assumed to follow a local linear trend model, corresponding to a random walk with a time-varying drift denoted here by δ_t, [17] we constraint s_t^C to follow a stationary AR(2) process. [18] As a whole, these assumptions are in line with the FEER-BEER approach. Indeed, in the case of the BEER model, we have by hypothesis: $(q_t, BEER_t) \sim CI(1,1)$, whereas the contribution of Barisone, Driver and Wren-Lewis (2006) shows exactly the same type of relation for the FEER, i.e. $(q_t, FEER_t) \sim CI(1,1)$, meaning that the implied misalignment is in both cases $I(0)$.

As recalled by Kim and Nelson (1999), the inclusion of unobserved factors in the specification of the model, requires the use of the state-space framework, made operational by the Kalman filter, in order to derive proper statistical inference in the time series context. More generally, state space modeling

16. For more details on the Kalman filter, see the Appendix C.1.2 or the seminal contribution of Kalman (1960).

17. See Chapter 2 in Durbin and Koopman (2001) for more details.

18. This last hypothesis is in line with the specification proposed by Engel and Kim (1999), which they chose to perfectly match their U.S./U.K. real exchange rate series. As it would have been time-consuming to specify an AR for each generated series and develop the associated state-space form of the full model accordingly, we consider here that all our series mimic the main patterns of that under scrutiny in Engel and Kim (1999) and thus, follow the same stochastic process for their cyclical component.

provides a unified methodology for treating a wide range of problems in time series analysis among which the Box and Jenkins ARIMA "toolbox" appears as a particular case. [19]

Consistently, the model can be re-written into a state-space form as follows:

$$s_t = \quad F_t\theta_t + v_t, \qquad v_t \sim \mathcal{N}_p(0, V_t) \tag{21}$$

$$\theta_t = \quad G_t\theta_{t-1} + w_t, \quad w_t \sim \mathcal{N}_m(0, W_t) \tag{22}$$

with:

$$F = \begin{pmatrix} 1 & 0 & 1 & 0 \end{pmatrix}, G = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi 2 & 0 \end{pmatrix}, V = (0), W = \begin{pmatrix} \sigma_\mu^2 & 0 & 0 & 0 \\ 0 & \sigma_\delta^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The vector of states being defined as follows: $\theta_t = \left(s_t^P, \delta_t, s_t^C, \theta_{t,4}\right)'$

2.2.2 A Bayesian analysis

The resulting dynamic linear model has five unknown parameters, namely σ_ε^2, σ_z^2 for the trend component, and ϕ_1, ϕ_2 and σ_u^2 for the AR(2) component that we summarize into the vector ψ. As recalled by Berger and Kempa (2010), two approaches can be opposed. [20] In a classical, i.e. frequentist, framework, ψ is estimated by maximum likelihood. Its estimated value is then assumed constant and incorporated into the state-space representation of the model so that the usual Kalman filter and smoother may apply. From a Bayesian stand-point, quoting Schorfheide (2008), 'prior information is used to down-weight the likelihood function in regions of the parameter space that are inconsistent with out-of-sample information and/or in which the structural model is not interpretable'. In our model, we assume that agents adopt this second approach. More specifically, we extend the usual trial-and-error learning strategies on which agents rely for the sequential choice of their forecasting rules to the way they model the fundamental value. Agents may have difficulties to perceive accurately the drivers of exchange rate fundamentals, which mislead them in the assessment of the medium-run equilibrium exchange rate, but they also can learn from their mistakes when new data become available over time. They are thus supposed to have a prior knowledge of the exchange rate fundamentals' drivers, summed up in ψ, that they update sequentially by observing the data over time according to a "learning process". We then treat ψ as a random quantity with a known prior density $p(\psi)$ and estimate the posterior density which turns out in this context to be the joint conditional distribution of the state vectors and the unknown parameters given the observations $p(\theta_{0:t}, \psi | y_{1:t})$, by combining information contained in $p(\psi)$ and the sample data $y_{1:t}$. Even if simple in principle, those calculations involve computations that are usually not analytically manageable. Thus, simulation methods such as Markov Chain Monte Carlo (MCMC) algorithms are implemented to approximate the posterior distribution of interest. [21]

19. For more details, see Chapter 3, pp. 46-52 from Durbin and Koopman (2001).
20. See Appendix C.1.3 for more insights on this point.
21. See Appendix C.2.1 for details,and Robert and Casella (2004) or Gamerman and Freitas Lopes (2006) for reviews on MCMC methods.

In practice, we follow Petris, Petrone and Campagnoli (2009) by assuming two sets of independent conjugate priors for $\psi_1 = (\phi_1, \phi_2)$ and $\psi_2 = (\sigma_\varepsilon^{-2}, \sigma_z^{-2}, \sigma_u^{-2})$ respectively, whose specifications (detailed *infra*) lead to a normal-gamma prior for ψ, which is a standard hypothesis in the context of a linear Gaussian model (see Geweke, 2005, pp. 40-41). [22]

Let us describe in details the assumed specifications for each sub-set of prior distributions.

Autoregressive parameters ψ_1: On the one hand, we assume the product of two normal distributions as a prior for the auto-regressive parameters ϕ_1 and ϕ_2 which gives:

$$(\phi_1, \phi_2) \sim \mathcal{N}\left(0, \left(\frac{2}{3}\right)^2\right) \times \mathcal{N}\left(0, \left(\frac{1}{3}\right)^2\right), \text{restricted to } \Omega \tag{23}$$

with Ω defining the stationarity region of the cyclical process s_t^g:

$$\Omega = \{\phi_1 + \phi_2 < 1, \phi_1 - \phi_2 > -1, |\phi_2| < 1\}$$

Precision parameters ψ_2: On the other hand, we set as priors for the precision parameters, i.e. the inverse of the variances, three independent gamma distributions defined as follows:

$$\sigma_\varepsilon^{-2}, \sigma_z^{-2}, \sigma_u^{-2} \sim \mathcal{G}\left(\frac{a^2}{b}, \frac{a}{b}\right) \tag{24}$$

with $\mathbf{E}\left(\sigma_i^{-2}\right) = a$ and $Var\left(\sigma_i^{-2}\right) = b, \forall i \in \{\varepsilon, z, u\}$ After calculations (see proof in Appendix C.2.1), we obtain the following conditionally independent and gamma-distributed precisions:

$$\sigma_\varepsilon^{-2}|\cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,1} - (G_t\theta_{t-1,1})_1\right)^2\right) \tag{25}$$

$$\sigma_z^{-2}|\cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,2} - (G_t\theta_{t-1,2})_2\right)^2\right) \tag{26}$$

$$\sigma_u^{-2}|\cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,3} - (G_t\theta_{t-1,3})_3\right)^2\right) \tag{27}$$

with: $\sigma_i^{-2}|\sigma_j^{-2}, \sigma_k^{-2}, \theta_{0:t}, y_{1:t} \sim \sigma_i^{-2}|\ldots, \forall (i,j,k) \in \{\varepsilon, z, u\}, i \neq j, i \neq k,$ and $j \neq k$

In this context of a separable unknown parameter vector, a hybrid sampler may be run. [23] It consists in the drawing of first ψ_1 given ψ_2 and the data $y_{1:T}$, second, of the states given ψ and $y_{1:T}$, and finally of ψ_2 given ψ_1 and the states. Whereas the first and three steps may be run by means either of direct sampling or MCMC methods in case of densities with unknown properties, the drawing of the states is straightforward and can be easily performed through the Forward Filtering Backward Sampling algorithm (FFBS). [24] As recalled by Petris et al. (2009), the AR parameters, given the precisions have a

22. As a first approach, we stick to conjugate priors, which present the advantage to provide a closed-form expression for the posterior, avoiding potential difficult numerical integration. For more details on the choice of priors and their properties, see Chapter 4 from Bauwens, Lubrano and Richard (1999), and more specifically regarding conjugate priors, Chapter 2, pp.38-46 from Geweke (2005).

23. See Appendix C.2.2 for technical details concerning the implementation of the hybrid sampler.

24. See Frühwirth-Schnatter (1994) for details.

non-standard distribution. Thus, a direct sampling approach cannot be implemented. As the full conditional distribution is difficult to sample from, a Metropolis-Hastings step will replace the corresponding Gibbs sampling method. More specifically, we implement Gilks, Best and Tan (1995)'s Adaptive rejection Metropolis sampler to draw samples of $\psi_1 = (\phi_1, \phi_2)'$. Turning to the precision parameters, due to their assumed gamma-distribution, they can be easily drawn from the usual Gibbs filter. All those steps are summed up in Figure 1 [25] and lead to the derivation of the full posterior conditional density associated to our state-space model.

Figure 1 about here

Once N samples of the parameters have been drawn from the posterior distribution, relying on the Law of Large Numbers, one can approximate any function $g(\psi)$ having a finite posterior expectation by the Monte Carlo ergodic mean as follows:

$$\mathbf{E}_p\left[g(\psi)\right] \approx N^{-1} \sum_{j=B+1}^{N} g\left(\psi_j\right) \tag{28}$$

where B represents the burn-in period, i.e. the number of iterations that are necessary to reach the limit distribution of interest p.

In practice, medium-run agents, have a prior idea of the vector of parameters ψ, they then assume it to be fixed over a certain period, and update their prior knowledge at a certain point of time, [26] using the observations according to a Bayesian rule. Depending on the assumed initial values, this correction may be more or less abrupt. In the simulation exercise that follows, we assume three heuristic cases and compare the implied dynamics for the exchange rate.

3 Simulation results

In what follows, we examine how the confrontation of the different forecasting rules impact FX rate dynamics. As recalled by De Grauwe and Grimaldi (2006*a*), the nonlinear structure of the model does not allow for simple analytical solutions. We thus have to perform numerical simulations in order to derive some key features. Each set of stochastic simulations consists here in 10000 points. They correspond to, first, the draw of 5000 points considering only two forecasting rules, namely the short and the long-run rules, and then, to the draw of 5000 other points once the medium-run rule has been introduced, this for two reasons. First, because the medium-run approach needs the past of the exchange rate series to be implemented and second, because this allows us to study the differences in the exchange rate dynamics before and after the adoption of the medium-run rule.

As initial values for the parameters of the behavioral finance model, we take those proposed in De Grauwe and Grimaldi (2006*a*) as a benchmark. [27] All the initial values are summed up in Table 3 that follows.

25. This representation is taken from Petris et al. (2009).

26. We assume here that there exists a certain inertia in agents' updates. More precisely, they do not perceive and react

Table 1: Initial values for our behavioral finance model

ψ	β	η_M	η_L	γ	r	r^*	s_0^*	s_0
0.2	0.9	0.2	0.2	5	0	0	0	0

Regarding the medium-run model, two sets of values must be defined. First, due to the prior structure we assume for ψ, we have to choose likely values for a and b, which characterize the prior opinion of agents about the unknown precisions, namely σ_ε^{-2}, σ_z^{-2}, and σ_u^{-2}, in terms of their mean and variance respectively. Here, we set $a = 1$ and $b = 1000$, which allows for a relatively high level of prior uncertainty. Second, we must determine initial values for ψ. We assume here three different cases depending on medium-run agents' ability to correctly perceive and update information they have access to.[28] More specifically, medium-run agents may be:

Scenario 1: right, in the sense that the initial values for the parameters (ψ) are equal to the Maximum Likelihood Estimates (MLE) based on the previous 5000 points;

Scenario 2: wrong, in the sense that the initial values for ψ are chosen to be far from the MLE estimates;

Scenario 3: wrong then right, trough a "learning" Bayesian process.

As mentioned before, we make the assumption that there exists a certain inertia in medium-run agents' Bayesian update process. This hypothesis means that agents do not perceive and react instantaneously to a potential inadequacy of their prior idea of ψ and the data. This may be due to the intrinsic nature of information they have access to. Indeed, as showed by Oberlechner and Hocking (2004), the pieces of information agents use to make decisions are not neutral in the sense that FX trading participants (the agents in our model) and financial news media are dependent on each other. More specifically, 'information of the news services often consists of trading participants' perceptions and interpretations of the market, which are fed back to the traders in the market. As a result, a highly circular cycle of collective information processing in the market emerges.' This circularity bounds the set of possible values to those forecast by the agents, and lead to a certain inertia in the revision process.

Moreover, agents may be unable to correctly predict the parameters. They may be wrong due to their inability to correctly forecast the fundamentals, or even to predict which fundamentals play a key role. For instance, Christopoulos and León-Ledesma (2009) show in a recent work, that multi-step causality analysis puts into light much more complex links between fundamentals and exchange rates than tested before. According to those authors, given the modern design of monetary policy, one would expect that the temporal link between exchange rates and fundamentals will depend on the timing of policy rules and expectations formation. As a result, a particular fundamental may not be useful to forecast the exchange

instantaneously to potential changes of the market characteristics. See Section 3 for more insights on this assumption.

27. In their approach, a sensitivity analysis is performed on two key parameters, namely β, the extrapolation parameter of short-run agents and γ, the sensitivity of agents to relative profitability. They show that as a whole, the rise in one or the other leads to significantly decrease the occurrence of stable fundamentalist-type *equilibria*. For more details, see De Grauwe and Grimaldi (2006*a*) pp.30-32.

28. See Appendix C.2.3 for full details on the assumed initial values for the set of parameters.

rate one period ahead but it may do so more than one period ahead if it helps forecasting a third variable which, in turn, can help predict exchange rates one period ahead. Finally, there may exist an inconsistency between the fundamental values agents can observe and consider in their prediction exercise, and those that effectively drive the exchange rate in real-time. For instance, Ehrmann and Fratzscher (2005) show that their model based on real-time data explains about 75% of the monthly directional changes of the US dollar-euro exchange rate.

Let us first draw two typical examples of simulated data derived from the confrontation of short and long-run agents.

Figures 2 and 3 about here

In each case, the current exchange rate diverges from its fundamental value only when all the agents adopt the short-run (i.e. chartist) rule. On the contrary, when agents switch frequently between the short and the long-run (i.e. fundamentalist) rules, the observed dynamics is stable in the sense that the current exchange rate stays around its long-run fundamental value. There is thus a clear opposition between chartists and fundamentalists in terms of stabilizing characteristics. Whereas long-run agents' actions drive the exchange rate closed to its fundamental value, the chartists reproduce the past dynamics, which can lead to potential long-lasting disconnections of the exchange rate from its long-run equilibrium. We replicate here one of the main results of De Grauwe and Grimaldi (2006*a*) model, namely, the reproduction in a very convincing way of what Obstfeld and Rogoff define as the "exchange rate disconnect puzzle".[29]

What happens now when we introduce a third player in the game, i.e. a medium-run approach, allowing agents to switch between three forecasting rules, namely the short, the medium and the long-run ones? As the medium-run rule is of a fundamentalist type, we expect it to be stabilizing, but is this characteristic conditioned by agents' good perception of the fundamentals' drivers? To give some insights on this issue, we propose to take a look at various outputs resulting from the three *scenarii* we have introduced earlier.

What happens when agents are right, i.e. when they rely on the MLE estimates for ψ?

Figure 4 about here

We can see in both cases that the introduction of the medium-run rule leads to a drastic stabilization of the current rate closed to its PPP value. This correction stems from the fact that the filtered value, i.e. the medium-run equilibrium exchange rate, is very closed to the current one. Mechanically, the medium-run negative feedback rule is almost the exact opposite to the short-run positive one, corresponding to a "leaning against the wind strategy". Medium-run agents are thus reinforcing the mean-reverting dynamics in the market, thereby strengthening the hand of long-run agents at the expense of the short-run ones, see Chapter 9 from De Grauwe and Grimaldi (2006*a*) or Westerhoff and Dieci (2006); Westerhoff

29. See Obstfeld and Rogoff (2000).

(2008) for more details.

What happens now when agents are wrong, i.e. when they rely on values far from the MLE estimates for ψ?

Figures 5 and 6 about here

We can see from Figure 5 that the medium-run rule seems to stay stabilizing in the first case, although to the cost of greater volatility. Considering other simulations, we may show that this stabilizing characteristic is purely random, see for instance Figures 7 related to two other simulations derived from the same set of parameters, that lead to a divergence of the exchange rate in spite of a prior convergence. In addition, when we look at the second case in Figure 5, i.e. when the exchange rate is already experiencing a long-lasting disconnection from its PPP value, we can see that the introduction of the medium-run rule cannot drive the current exchange rate towards its long-run fundamental rate. Since important switches can be observed from Figure 6 between short and medium-run agents, this disconnection is due to both types. As whole, we conclude from those two sets of simulations that the stabilizing properties of the medium-run rule seems to depend on agents' good perception of the fundamentals, summed up in ψ.[30]

Figure 7 about here

As a last exercise, and to push further our analysis, we assume now that agents may update their perception of the fundamentals' drivers ψ through a Bayesian "learning process". In other words, we ask the following question: What happens when agents are wrong but then learn from the data and update their knowledge in a Bayesian manner?[31] We can see from the graphs in Figure 8 (derived from the same first two sets of simulations), that medium-run agents remain destabilizing until the revision step, in time $t = 7500$. After incorporating truncated information from the data,[32] the parameters become mechanically far much closer to the MLE estimates and allow to go back to "leaning against the wind strategy" as in the previous case.

Figure 8 about here

4 Some preliminary analysis

Our ambition in this paper is twofold. First, we propose an original interpretation of the behavioral finance model applied to the exchange rate issue by including a reflexion in terms of forecasting time horizons that is consistent with recent survey analysis. According to this literature, instead of the usual confrontation between stabilizing "fundamentalists" vs. destabilizing "chartists", we can oppose agents

30. Other simulation sets confirm those conclusions. They are available upon request.

31. Note that this idea of misperception and "learning process" has also been explored by Lewis and Markiewicz (2009) in an empirical exercise, where heterogeneous agents rely on monetary models in the vein of Frenkel (1976) and Bilson (1978) for their forecasts.

32. In each case, Bayesian estimations have been performed on a sub-sample of 100 points.

who act on a daily basis and those who manage funds over a longer period, typically several weeks or months. We then disentangle three forecasting schemes related to the short, the long and, in between, the medium run. Whereas the first two rules stick to the usual representation, we incorporate an intermediary regime, of a fundamental type, that relies on the assessment of a medium-run equilibrium value for the exchange rate supposed to potentially differ from the PPP and assessed here through an APEER model, namely a trend-cycle decomposition as proposed by Engel and Kim (1999). Second, we extend the usual trial-and-error learning strategies on which agents rely for the sequential choice of their forecasting rules to the way they model the fundamental value. We assume that they may have difficulties to perceive accurately the drivers of exchange rate fundamentals, either because they do not know which fundamentals play a key role or because the data they have access to are not the ones that effectively drive the exchange rate in real-time. Consequently, agents may sometimes make mistakes and correct their knowledge with a certain delay. We have chosen here a Bayesian framework to tackle this issue, considering that agents may have a prior idea of what should be the drivers of the fundamentals but may also incorporate information from the data overtime, proceeding to sequential updates of their knowledge.

We show that this modeling leads to two major conclusions:

- We observe an opposition between stabilizing agents and destabilizing agents. In addition to the usual "battle" between stabilizing long-run (fundamentalist) vs. destabilizing short-run (chartist) agents, we show that medium-run agents may switch from one type to another depending on their ability to correctly perceive the drivers of exchange rate fundamentals. This means that even a fundamentalist type rule may lead to long-lasting disconnection between the exchange rate and its fundamental value;
- A good knowledge or perception of the fundamental drivers of the exchange rate is a precondition to make the median rule stabilizing. More specifically, in a situation of a current exchange rate disconnection from its fundamental value, medium-run agents are stabilizing only when they correctly predict ψ. Indeed, under this condition, the medium-run fundamental value is very closed to the current one, which makes this rule corresponding to a "leaning against the wind strategy".

As a whole, this piece of work departs from the traditional behavioral finance framework by putting lights on two major concerns, namely the reflexion about time horizons and the extension of agents' learning process to their perception of fundamentals. We do believe that the Bayesian framework offers a comprehensive approach to deal with those issues and may be extended in four different ways:

- To properly take into account the variety of time horizons agents deal with, one needs to re-think the evaluation mechanism they put into practice to assess the *ex-post* performances of their forecasting rules over time. Indeed, the fitness criterion applied here relies on a short-run view, since it is implemented at each step or period t, the γ parameter allowing for a certain inertia in the switching process. If the exchange rate anchor varies according to the time horizon (short, medium, or long run), then it will be expected that the chartist rule will outperform fundamentalist-type ones at least over a certain period. The key question seems then to explain why relying on longer time horizons may be profitable or to say it differently, why, while chartist rules are efficient, other agents chose to fix to other types, supposed to hold in the long run. It seems that underlying wealth of FX dealers and fund managers differs in size and time constraint in the sense that whereas the former must reach a neutral position on a daily basis, the latter experience more flexibility in the way they invest their assets (see Gehrig and Menkhoff, 2005 for more insights). This distinction may then be key to solve this issue. For instance, modifying the profit rule associated to each horizon or endogenizing γ by modeling explicitly agents' wealth could constitute potential extensions. In such frameworks, different market pressures in the short, medium or long run could be accounted for and provide an interesting complement to the analysis addressed in this paper.

- The stabilizing characteristic of the medium-run rule may depend on the structure of the prior we chose for the vector of unknown parameters ψ. As a robustness check, and as a way to describe a larger variety of potential situations for agents' prior knowledge, it would be interesting to consider different types of priors: What if we implement a flat, i.e. non-informative prior, expressing a situation of "prior ignorance", or another set of conjugate one, relaxing for example the stationarity constraint? Other types such as Jeffrey priors, also known as reference priors from an informational decision based may be explored.

- Another limitation comes with the idea that agents are unable to update their knowledge at each step of the simulation exercise. Even if this hypothesis may be explained by psycho-economic arguments, one could expect that those updates would take place at a higher frequency, which cannot be, due to great computational limitations. The Bayesian framework we rely on in this exercise is considerably time-consuming especially in the context of sequential models such as ours since any time a new observation becomes available, a totally new Markov Chain has to be simulated. To cope with this issue, sequential MCMC methods [33] that have proved to be successful in online filtering application of dynamic linear models with unknown parameters, are to be performed. [34]

- Finally, we could rely on a more advanced modeling for the assessment of the medium-run equilibrium exchange rate value. More specifically, as proposed by Berger and Kempa (2010), we could derive a state-space representation and apply the same estimation methodology to the recent Taylor-rule based monetary exchange rate model of Engel and West (2006) that has proved to perform fairly well at short and medium-run time horizons, see Molodtsova and Papell (2009) among others.

We leave all those developments for further research.

5 Conclusion

To conclude, this piece of work proposes a Bayesian extension of the traditional behavioral finance model "à la" De Grauwe and Grimaldi (2006*a*) in which agents differ in their forecasting time horizons as a way to re-explore Obstfeld and Rogoff's famous "exchange rate disconnect puzzle". As previously mentioned, those differences translate into the equilibrium values those agents target for the exchange rate. We then assume three rules corresponding to the short, the medium and the long run respectively. Whereas the first two rules stick to the usual representation, i.e. chartist vs. fundamentalist agents, we

33. See Cappé, Godsill and Moulines (2007) for a overview of the field and more specifically, Liu and West (2001) who propose an extended auxiliary particle filter approach devoted to the sequential estimation of dynamic linear models with unknown parameters.

34. See Chapter 5 of Petris et al. (2009) for an introductory presentation of this new branch.

incorporate an intermediary regime, of a fundamental type, that relies on the assessment of a medium-run equilibrium value for the exchange rate supposed to potentially differ from the PPP and assessed here through an APEER model. In addition, we further extend the usual trial-and-error learning strategies on which agents rely for the sequential choice of their forecasting rules to the way they model the medium-run fundamental value, through a Bayesian sequential updating procedure. This modeling leads to two major conclusions: (i) In addition to the usual 'battle' between stabilizing long-run (fundamentalist) vs. destabilizing short-run (chartist) agents, we show that medium-run agents may switch from one type to another depending on their ability to correctly perceive the drivers of exchange rate fundamentals. This means that even a fundamentalist type rule may lead to long-lasting disconnections between the exchange rate and its fundamental value; (ii) A good knowledge or perception of the fundamental drivers of the exchange rate is a precondition to make the median rule stabilizing. Further extensions discussed in the previous section are now to be done in order to test the robustness of those results in a more general context.

References

Barisone, G., Driver, R. and Wren-Lewis, S. (2006), 'Are Our FEERs Justified?', *Journal of International Money and Finance* **25**(5), 741–759.

Bask, M. (2007), 'Chartism and Exchange Rate Volatility', *International Journal of Finance and Economics* **12**(3), 301–316.

Bauwens, L., Lubrano, M. and Richard, J. (1999), *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, USA.

Berger, J. and Wolpert, R. (1984), *The Likelihood Principle*, Lectures notes - Monograph series, Institute of Mathematical Statistics.

Berger, T. and Kempa, B. (2010), Taylor Rules and the Canadian-US Equilibrium Exchange Rate, Working Papers 10/643, Ghent University, Faculty of Economics and Business Administration.

Beveridge, S. and Nelson, C. R. (1981), 'A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'', *Journal of Monetary Economics* **7**(2), 151–174.

Bilson, J. F. (1978), 'The Monetary Approach to the Exchange Rate: Some Empirical Evidence', *IMF Staff Papers* **25**, 48–75.

Bénassy-Quéré, A., Béreau, S. and Mignon, V. (2009), 'Robust Estimations of Equilibrium Exchange Rates: A Panel BEER Approach', *Scottish Journal of Political Economy* **56**(5), 608–633.

Bénassy-Quéré, A., Béreau, S. and Mignon, V. (2010), 'On the Complementarity of Equilibrium Exchange Rate Approaches', *Review of International Economics,* forthcoming .

Brock, W. A. and Hommes, C. H. (1997), 'A Rational Route to Randomness', *Econometrica* **65**(5), 1059–1096.

Cappé, O., Godsill, S. and Moulines, E. (2007), 'An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo', *Proceedings of the IEEE* **95**(5), 899–924.

Cheung, Y.-W., Chinn, M. D. and Marsh, I. W. (2004), 'How Do UK-based Foreign Exchange Dealers Think Their Market Operates?', *International Journal of Finance and Economics* **9**(4), 289–306.

Christopoulos, D. and León-Ledesma, M. A. (2009), On causal Relationships Between Exchange Rates and Fundamentals: Better Than You Think, Studies in Economics 0909, Department of Economics, University of Kent.

Clarida, R. and Gali, J. (1994), 'Sources of Real Exchange-Rate Fluctuations: How Important Are Nominal Shocks?', *Carnegie-Rochester Conference Series on Public Policy* **41**(1), 1–56.

Clark, P. and MacDonald, R. (1998), Exchange Rates and Economic Fundamentals: A Methodological Comparison of BEERs and FEERs, IMF Working Papers 98/67, International Monetary Fund.

Clark, P. and MacDonald, R. R. (2004), 'Filtering the BEER: a permanent and transitory decomposition', *Global Finance Journal* **15**(1), 29–56.

Cumby, R. E. and Huizinga, J. (1990), The Predictability of Real Exchange Rate Changes in the Short and Long Run, NBER Working Papers 3468, National Bureau of Economic Research, Inc.

De Grauwe, P. and Grimaldi, M. (2006*a*), *The Exchange Rate in a Behavioral Finance Framework*, Princeton University Press.

De Grauwe, P. and Grimaldi, M. (2006*b*), 'Exchange Rate Puzzles: A Tale of Switching Attractors', *European Economic Review* **50**(1), 1–33.

De Grauwe, P. and Rovira Kaltwasser, P. (2007), Modeling Optimism and Pessimism in the Foreign Exchange Market, CESifo Working Paper Series 1692, CESifo Group Munich.

Dornbusch, R. (1976), 'Expectations and Exchange Rate Dynamics', *The Journal of Political Economy* **84**(6), 1161–1176.

Driver, R. and Westaway, P. F. (2004), Concepts of Equilibrium Exchange Rates, Working Paper 248, Bank of England.

Durbin, J. and Koopman, S. (2001), *Time Series Analysis by State Space Methods*, Oxford University Press.

Ehrmann, M. and Fratzscher, M. (2005), 'Exchange Rates and Fundamentals: New Evidence from Real-time Data', *Journal of International Money and Finance* **24**(2), 317–341.

Engel, C. and Kim, C.-J. (1999), 'The Long-Run U.S./U.K. Real Exchange Rate', *Journal of Money, Credit and Banking* **31**(3), 335–56.

Engel, C. and West, K. D. (2006), 'Taylor Rules and the Deutschmark: Dollar Real Exchange Rate', *Journal of Money, Credit and Banking* **38**(5), 1175–1194.

Frankel, J. (1979), 'On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials', *The American Economic Review* **69**(4), 610–622.

Frankel, J. A. and Froot, K. A. (1990), 'Chartists, Fundamentalists, and Trading in the Foreign Exchange Market', *American Economic Review* **80**(2), 181–85.

Frenkel, J. (1976), 'A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence', *The Scandinavian Journal of Economics* **78**(2), 200–224.

Frühwirth-Schnatter, S. (1994), 'Data Augmentation and Dynamic Linear Model', *Journal of Time Series Analysis* **15**(2), 183–202.

Froot, K. A. and Rogoff, K. (1995), Perspectives on PPP and Long-run Real Exchange Rates, *in* G. M. Grossman and K. Rogoff, eds, 'Handbook of International Economics', Elsevier.

Gamerman, D. and Freitas Lopes, H. (2006), *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Texts in Statistical Science Series, 2nd edn, Chapman and Hall.

Gehrig, T. and Menkhoff, L. (2003), Technical Analysis in Foreign Exchange - The Workhorse Gains Further Ground, Diskussionspapiere dp-278, Universität Hannover, Wirtschaftswissenschaftliche Fakultät.

Gehrig, T. and Menkhoff, L. (2004), 'The Use of Flow Analysis in Foreign Exchange: Exploratory Evidence', *Journal of International Money and Finance* **23**(4), 573–594.

Gehrig, T. and Menkhoff, L. (2005), 'The Rise of Fund Managers in Foreign Exchange: Will Fundamentals Ultimately Dominate?', *The World Economy* **28**(4), 519–540.

Geweke, J. (1989), 'Bayesian inference in econometric models using monte carlo integration', *Econometrica* **57**(6), 1317–39.

Geweke, J. (2005), *Contemporary Bayesian Econometrics and Statistics*, Wiley-Interscience.

Gilks, W., Best, N. and Tan, K. (1995), 'Adaptive Rejection Metropolis Sampling within Gibbs Sampling', *Applied Statistics* **44**(4), 455–472.

Gonzalo, J. and Granger, C. W. J. (1995), 'Estimation of Common Long-Memory Components in Cointegrated Systems', *Journal of Business and Economic Statistics* **13**(1), 27–35.

Hommes, C. (2008), Interacting Agents in Finance, *in* S. N. Durlauf and L. E. Blume, eds, 'The New Palgrave Dictionary of Economics', Palgrave Macmillan, Basingstoke.

Hooper, P. and Morton, J. (1982), 'Fluctuations in the Dollar: A Model of Nominal and Real Exchange Rate Determination', *Journal of International Money and Finance* **1**, 39–56.

Huizinga, J. (1987), 'An Empirical Investigation of the Long-run Behavior of Real Exchange Rates', *Carnegie-Rochester Conference Series on Public Policy* **27**(1), 149–214.

Kalman, R. E. (1960), 'A New Appraoch to Linear Filtering and Prediction Problems', *Transactions of the ASME–Journal of Basic Engineering* **82**(Series D), 35–45.

Kim, C.-J. and Nelson, C. R. (1999), *State-Space Models with Regime Switching*, The MIT Press.

Lewis, V. and Markiewicz, A. (2009), ‘Model Misspecification, Learning and the Exchange Rate Disconnect Puzzle’, *The B.E. Journal of Macroeconomics* **9**(1), Art. 13.

Liu, J. and West, M. (2001), Combined Parameter and State Estimation in Simulation-Based Filtering, *in* A. Doucet, N. de Freitas and N. Gordon, eds, ‘Sequential Monte Carlo Methods in Practice’, Springer, pp. 197–223.

Lyons, R. (2001), *The Microstructure Approach to Exchange Rates*, MIT Press.

MacDonald, R. R. (2000), ‘Concepts to Calculate Equilibrium Exchange Rates: An Overview’, *Discussion Paper, Economic Research of the Deutsche Bundesbank* **3**, 1–70.

Mark, N. C. (1995), ‘Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability’, *American Economic Review* **85**(1), 201–18.

Mark, N. C. and Sul, D. (2001), ‘Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel’, *Journal of International Economics* **53**(1), 29–52.

Meese, R. and Rogoff, K. (1983), ‘Empirical Exchange Rate Models in the Seventies: Do they Fit Out of Sample?’, *Journal of International Economics* **14**(1-2), 3–24.

Molodtsova, T. and Papell, D. H. (2009), ‘Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals’, *Journal of International Economics* **77**(2), 167–180.

Neely, C. J. (1997), ‘Technical Analysis in the Foreign Exchange Market: A Layman’s Guide’, *Federal Reserve Bank of Saint Louis Review* **79**(1), 23–38.

Oberlechner, T. and Hocking, S. (2004), ‘Information Sources, News, and Rumors in Financial Markets: Insights into the Foreign Exchange Market’, *Journal of Economic Psychology* **25**(3), 407–424.

Obstfeld, M. and Rogoff, K. (1995), ‘Exchange Rate Dynamics Redux’, *Journal of Political Economy* **103**(3), 624–60.

Obstfeld, M. and Rogoff, K. (2000), The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?, NBER Working Papers 7777, National Bureau of Economic Research, Inc.

Petris, G., Petrone, S. and Campagnoli, P. (2009), *Dynamic Linear Models with R*, Springer.

Piccillo, G. (2009), Exchange Rates and Asset Prices: Heterogeneous Agents at Work, Working Paper Series 09.03, Centre for Economic Studies, Katholieke Universiteit Leuven.

Robert, C. and Casella, G. (2004), *Monte Carlo Statistical Methods*, Springer, 2nd ed.

Rovira Kaltwasser, P. (2010), ‘Uncertainty about Fundamentals and Herding Behavior in the FOREX Market’, *Physica A* **389**(6), 1215–1222.

Schorfheide, F. (2008), Bayesian methods in macroeconometrics, *in* S. N. Durlauf and L. E. Blume, eds, ‘The New Palgrave Dictionary of Economics’, Palgrave Macmillan, Basingstoke.

Taylor, M. and Allen, H. (1992), ‘The Use of Technical Analysis in the Foreign Exchange Market’, *Journal of International Money and Finance* **11**(3), 304–314.

Westerhoff, F. H. (2008), 'The Use of Agent-Based Financial Market Models to Test the Effectiveness of Regulatory Policies', *Journal of Economics and Statistics* **228**(2), 195–227.

Westerhoff, F. H. and Dieci, R. (2006), 'The Effectiveness of Keynes-Tobin Transaction Taxes when Heterogeneous Agents can Trade in Different Markets: A Behavioral Finance Approach', *Journal of Economic Dynamics and Control* **30**(2), 293–322.

Williamson, J. (1985), *The Exchange Rate System*, second edn, MIT Press.

A Figures

Figure 1: Procedure of the Hybrid Sampler

1. **Initialization phase: we set** $\psi_2 = \psi_2^{(0)}$
2. **For** $i = 1, \ldots, N$**:**
   (a) **Draw** $\psi_1^{(i)}$ **from** $p(\psi_i | y_{1:T}, \psi_2 = \psi_2^{(i-1)})$ **using an Adaptive Rejection Metropolis-Hasting Sampler (ARMS)**
   (b) **Draw** $\theta_{0:T}^{(i)}$ **from** $p(\theta_{0:T} | y_{1:T}, \psi_1 = \psi_1^{i}, \psi_2 = \psi_2^{(i-1)})$ **using a Forward Filtering Backward Sampler (FFBS)**
   (c) **Draw** $\psi_2^{(i)}$ **from** $p(\psi_2 | y_{1:T}, \theta_{0:T}, \psi_1 = \psi_1^{(i)})$ **using a Gibbs Sampler**

Figure 2: Simulated data - convergence case

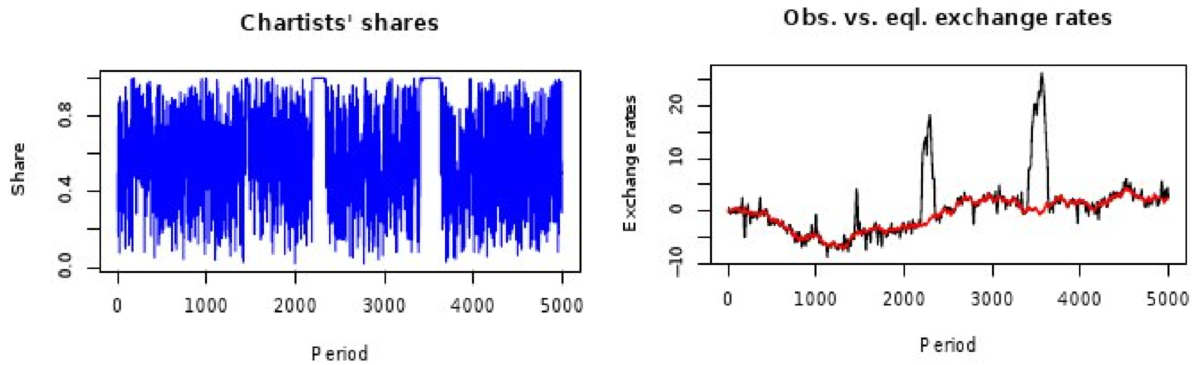


Figure 3: Simulated data - divergence case

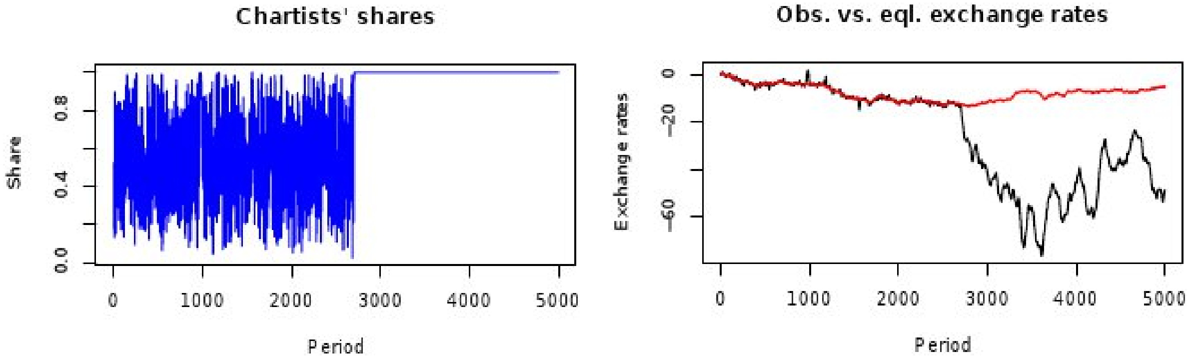
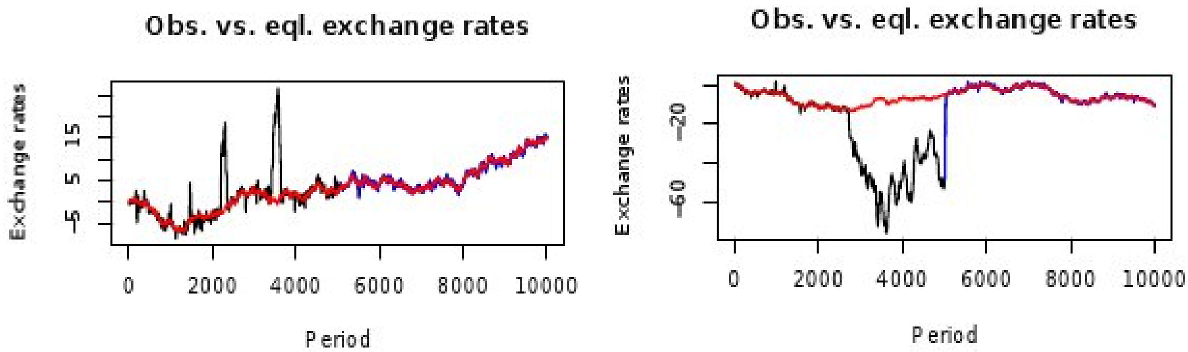


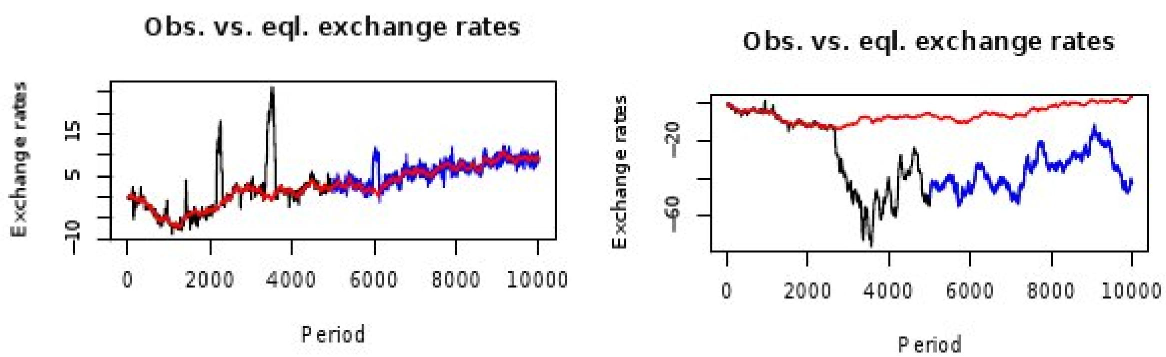
Figure 4: What happens when medium-run agents are right?



Convergence case

Divergence case

Figure 5: What happens when medium-run agents are wrong?



Convergence case

Divergence case

Figure 6: Who is responsible for this divergence?

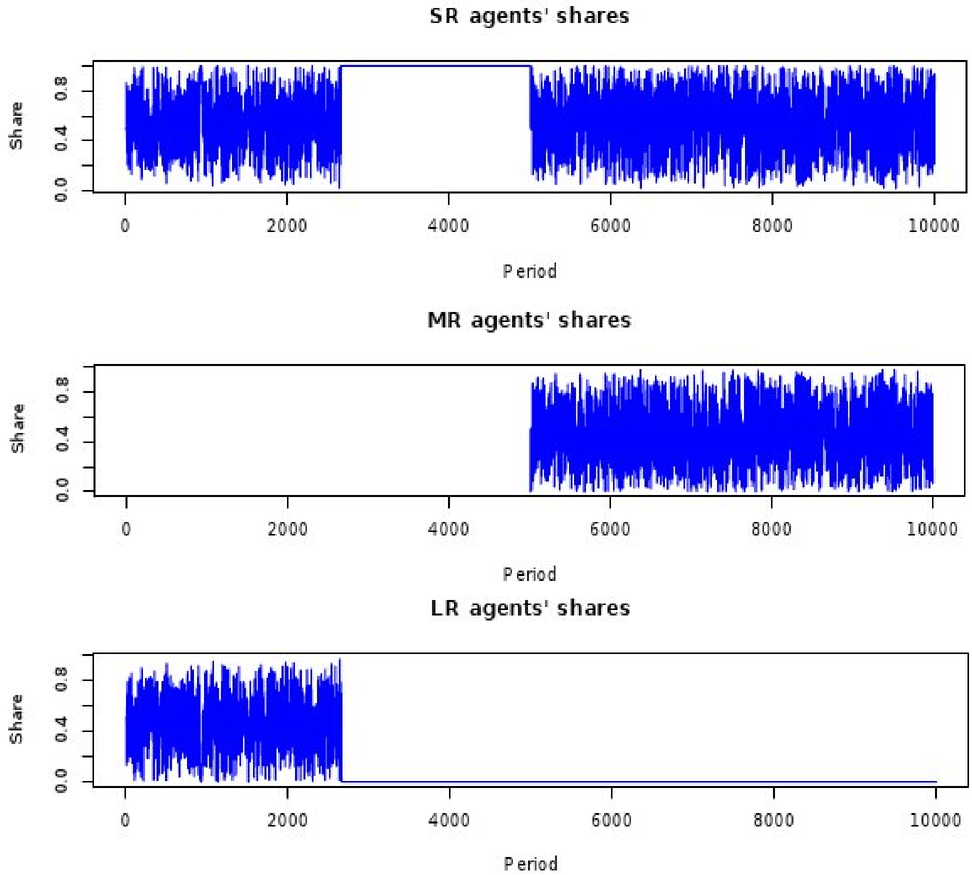


Figure 7: Other cases of divergence with prior convergence

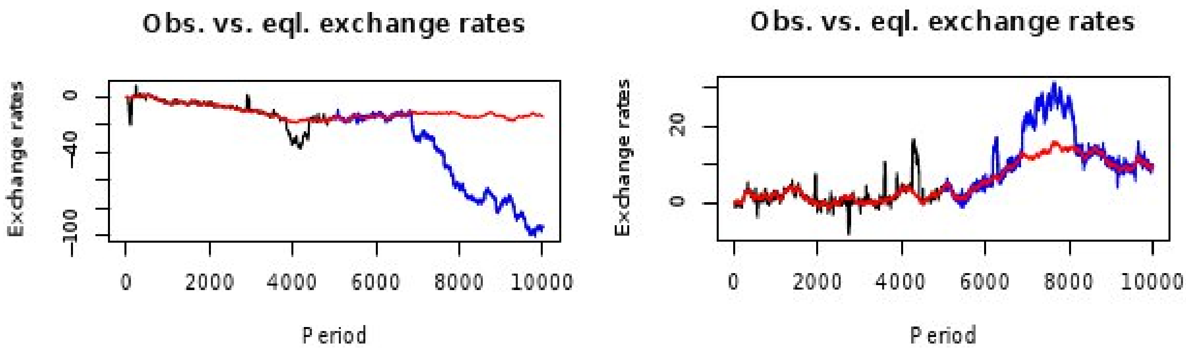
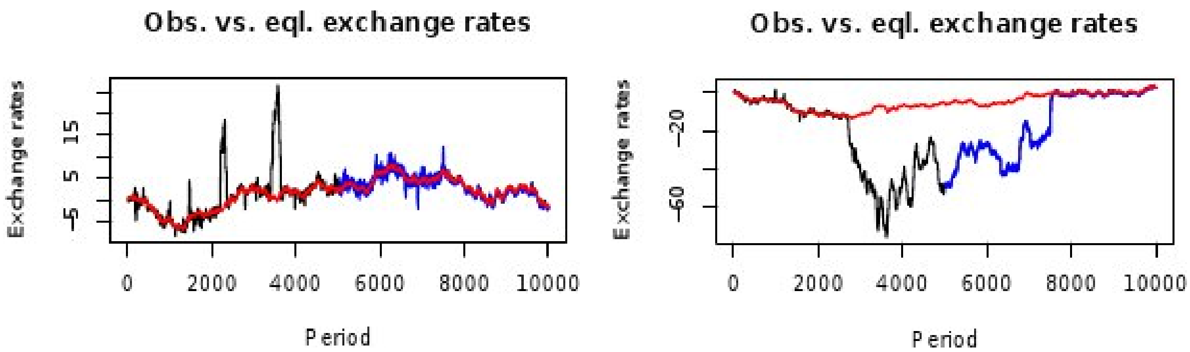


Figure 8: What happens when medium-run agents are wrong but learn from the data?



Convergence case

Divergence case

B A less simple presentation of the behavioral finance framework

We recall here a 'slightly more complex behavioral finance model' as presented in Chapter 3 of De Grauwe and Grimaldi (2006*a*). As mentioned by the authors, the model consists in three building blocks. First, agents decide the optimal portfolio using a mean-variance utility framework. Second, they make forecasts about the future exchange rate based on simple rules. Third, they evaluate their rules *ex-post* by comparing their relative risk-adjusted profitability by means of a fitness criterion proposed by Brock and Hommes (1997). Whereas steps 2 and 3 correspond to those already presented in sub-section 2.1 (from Equations 1 to 8), step 1 replaces the *ad hoc* specification of the future exchange rate as the weighted sum of predicted exchange rates for each type of agents, as specified in Equations 9 and 10. We are thus going to focus on the modeling of agents' optimal portfolio choice in the following lines.

Each agent invests in two assets, a domestic risk-free asset and a foreign risky one. The agents' utility function is specified as follows:

$$\mathcal{U}(W_{i,t+1}) = \mathbf{E}_t(W_{i,t+1}) - \frac{1}{2}\mu V_i(W_{i,t+1}), \forall i \in \{S, M, L\} \tag{29}$$

where $W_{i,t+1}$ represents the wealth of agents of type i at time $(t+1)$ (expressed in domestic currency), $V_i(W_{i,t+1})$ the conditional variance of agents of type i's wealth and μ the risk aversion as previously defined in Equation (6).

The wealth of agent i is characterized as follows:

$$W_{i,t+1} = (1+r^*)s_{t+1}d_{i,t} + (1+r)(W_{i,t} - s_t d_{i,t}) \tag{30}$$

with $d_{i,t}$ the holdings in foreign assets by agents of type i at time t, and r, r^* the domestic and foreign interest rates as defined in Equation (8). Thus, the first term of the RHS of the above expression represents the value of the (risky) foreign portfolio expressed in domestic currency whereas the second term corresponds to the value of the (riskless) domestic portfolio, both of them being assessed at time $(t+1)$. The substitution of Equation (29) into (30) and the resolution to the maximization program with respect to $d_{i,t}$ allows us to derive the following optimal condition:

$$d^o_{i,t} = \frac{(1+r^*)\mathbf{E}_{i,t}(s_{t+1}) - (1+r)s_t}{\mu\sigma^2_{i,t}} \tag{31}$$

with $\sigma^2_{i,t} = (1+r^*)^2 V_{i,t}(s_{t+1})$. The optimal holding of the foreign asset, $d^o_{i,t}$, depends on the expected excess return (corrected for risk) of the foreign asset.
The market demand for foreign assets at time t is the sum of individual demands, which leads to:

$$\sum_{i \in \{S,M,L\}} n_{i,t} d^o_{i,t} = D_t \tag{32}$$

where $n_{i,t}$ is the number of agents of type i.

Market equilibrium implies that the market demand is equal to the market supply, Z_t, assumed to be exogenous,[35] which leads to:

$$Z_t = D_t \tag{33}$$

Substituting the optimal holdings into the market demand and then into the market equilibrium equation for the exchange rate s_t yield the market clearing exchange rate:

$$s_t = \left(\frac{1+r^*}{1+r}\right) \frac{1}{\sum_{i\in\{S,M,L\}} \frac{\omega_{i,t}}{\sigma^2_{i,t}}} \left\{ \sum_{i\in\{S,M,L\}} \omega_{i,t} \frac{\mathbf{E}_{i,t}(s_{t+1})}{\sigma^2_{i,t}} - \Omega_t Z_t \right\} \tag{34}$$

with $\omega_{i,t}$ the share of agents of type i defined as follows:

$$\omega_{i,t} = \frac{n_{i,t}}{\sum_{i\in\{S,M,L\}} n_{i,t}} \tag{35}$$

and:

$$\Omega_t = \frac{\mu}{(1+r^*) \sum_{i\in\{S,M,L\}} n_{i,t}} \tag{36}$$

As in our model, the exchange rate is thus determined by the weighted expectations of agents about the future exchange rate. What differs here is that their forecasts are weighted by their respective variances $\sigma^2_{i,t}$. When agents of type i's forecasts have a high variance, the weight of those agents in the determination of the market exchange rate is reduced.

C State-space model representation of the medium run model

C.1 The general case[36]

C.1.1 Definition

A state-space model consists in a couple $(\theta_t, Y_t) \in \{\mathbf{R}^p \times \mathbf{R}^m\}$ with θ_t the state vector and Y_t the observation data, that satisfy the two following assumptions:

- The process $\{\theta_t\}$ is a Markov chain:

$$p(\theta_t|\theta_{0:t-1}, y_{1:t-1}) = p(\theta_t|\theta_{t-1}), \forall t \tag{37}$$

- Conditionally on $\{\theta_t\}$, the Y_t's are independent and Y_t only depend on θ_t:

$$p(y_t|\theta_{0:t-1}, y_{1,t-1}) = p(y_t|\theta_t), \forall t \tag{38}$$

This leads to the following condition for all t:

$$p(\theta_{0:t}, y_{1:t}) = p(\theta_0) \cdot \prod_{j=1}^{t} p(\theta_j|\theta_{j-1}) \cdot p(y_j|\theta_j) \tag{39}$$

35. The authors relax this assumption in further developments: they first model the current account dynamics in Chapter 4 and then study Central Bank interventions and their impact on Z_t in Chapter 8, see De Grauwe and Grimaldi (2006*a*) for further details.

36. Those developments are built on elements taken from Durbin and Koopman (2001) and Petris et al. (2009).

C.1.2 The Gaussian linear model and the derivation of the Kalman filter

An important class of state-space models is given by Gaussian linear state-space models also called *Dynamic Linear Models* (DLM). They are characterized by the following equations:

$$\text{Prior Gaussian distribution of the state-vector at time } t=0\text{: } \theta_0 \sim \mathcal{N}_p(m_0, C_0) \tag{40}$$

$$\begin{aligned} \text{Observation equation: } Y_t &= F_t\theta_t + v_t, \quad v_t \sim \mathcal{N}_p(0, V_t) && (41)\\ \text{State equation: } \theta_t &= G_t\theta_{t-1} + w_t, \quad w_t \sim \mathcal{N}_p(0, W_t) && (42) \end{aligned}$$

where F_t and G_t are assumed to be known matrices of order (m, p) and (p, p) respectively, and $(v_t)_{t\geq 1}$ and $(w_t)_{t\geq 1}$ are two independent sequences of independent Gaussian random vectors with mean 0 and known variance matrices V_t and W_t respectively. It is furthermore assumed that θ_0 is independent of $\{v_t\}$ or $\{w_t\}$. It can be shown that under those conditions, the DLM satisfies equations (37) and (38) with $Y_t|\theta_t \sim \mathcal{N}(F_t\theta_t, V_t)$ and $\theta_t|\theta_{t-1} \sim \mathcal{N}(G_t\theta_{t-1}, W_t)$.

We assume as a first step that the densities $p(y_t|\theta_t)$ and $p(\theta_t|\theta_{t-1})$ are fully specified. In practice, they depend of an unknown vector of parameters (defined as ψ in the core of the paper) that we may estimate using two different approaches (see sub-section C.1.3). To estimate the state-vector we compute the conditional densities $p(\theta_s|y_{1:t})$. Whether s is equal, lower or *greater* than t we deal with the *filtering*, *smoothing*, and *forecasting* problems. We focus here on the derivation of the Kalman filter that medium-run agents are assumed to implement to derive their equilibrium value for the exchange rate.

Assuming the previous description for the DLM, and that $\theta_{t-1}|y_{1:t-1} \sim \mathcal{N}(m_{t-1}, C_{t-1})$, we get:

1. The one-step-ahead predictive distribution of θ_t given $y_{1:t-1}$:

$$\begin{aligned} \theta_t|y_{1:t-1} \sim \mathcal{N}(a_t, R_t), \text{ with: } a_t &= \mathbf{E}(\theta_t|y_{1:t-1}) = G_t m_{t-1}\\ R_t &= Var(\theta_t|y_{1:t-1}) = G_t C_{t-1} G_t' + W_t \end{aligned}$$

2. The one-step-ahead predictive distribution of Y_t given $y_{1:t-1}$:

$$\begin{aligned} Y_t|y_{1:t-1} \sim \mathcal{N}(f_t, Q_t), \text{ with: } f_t &= \mathbf{E}(Y_t|y_{1:t-1}) = F_t a_t\\ Q_t &= Var(Y_t|y_{1:t-1}) = F_t R_t F_t' + V_t \end{aligned}$$

3. And finally, the filtering distribution of θ_t given $y_{1:t}$:

$$\begin{aligned} \theta_t|y_{1:t} \sim \mathcal{N}(m_t, C_t), \text{ with: } m_t &= \mathbf{E}(\theta_t|y_{1:t}) = a_t + R_t F_t' Q_t^{-1}(Y_t - f_t)\\ C_t &= Var(\theta_t|y_{1:t}) = R_t - R_t F_t' Q_t^{-1} F_t R_t \end{aligned}$$

Lemna 1 Law of Iterated Expectations
The Law of Iterated Expectations stands that:

$$\begin{aligned} \mathbf{E}\,(y|x) &= \mathbf{E}\big[\mathbf{E}\,(y|x,z)\,|x\big] && (43)\\ Var\,(y|x) &= \mathbf{E}\big[Var\,(y|x,z)\,|x\big] + Var\big[\mathbf{E}\,(y|x,z)\,|x\big] && (44) \end{aligned}$$

Lemna 2 Bayes Theorem
In the discrete case, the Bayes Theorem relates the conditional and marginal probabilities of events A and B, where B has a non-vanishing probability as follows:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)} \tag{45}$$

with $\mathbb{P}(A|B)$ *the conditional probability of A given B or posterior probability,* $\mathbb{P}(B|A)$ *the conditional probability of B given A, or the likelihood,* $\mathbb{P}(A)$ *the marginal probability of A, called the prior, and* $\mathbb{P}(B)$ *the marginal probability of B, acting as a normalizing constant.*

Lemna 3 Block Matrix Inversion
Let A, C be regular symmetric matrices of dimension (n, n) *and* (m, m) *respectively, and B a matrix of dimension* (n, m)*. We then get the following result:*

$$(A^{-1} + BC^{-1}B')^{-1} = A - AB(C + B'AB)^{-1}B'A \tag{46}$$

Proposition 1 Kernel of a Normal distribution
Assuming that $Y \sim \mathcal{N}(m, \Sigma)$*, the density of* Y*,* $p(y)$*, is given by:*

$$\begin{aligned} p(y) &= (2\pi)^{-\frac{1}{2}} |\Sigma|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(y-m)' \Sigma^{-1}(y-m)\right\} \\ &= (2\pi)^{-\frac{1}{2}} |\Sigma|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\Big(y'\Sigma^{-1}y - 2y'\Sigma^{-1}m + m'\Sigma^{-1}m\Big)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\Big(y'\Sigma^{-1}y - 2y'\Sigma^{-1}m\Big)\right\} \end{aligned}$$

We now exhibit the sketch of proof for the derivation of the Kalman filter in the case of a DLM.

Proof. The random vector $(\theta_0, \theta_1, \ldots, \theta_t, Y_1, \ldots, Y_t)$ has joint distribution given by Equation (39), where the marginal and conditional distributions involved are Gaussian. It follows that the joint distribution $(\theta_0, \theta_1, \ldots, \theta_t, Y_1, \ldots, Y_t)$ is Gaussian for any $t \geq 1$. Consequently, the distribution of any sub-vector including the conditional distribution of some components given other components are also Gaussian. Therefore, the predictive and filtering distributions are Gaussian. Their mean and variances can be derived as follows:

1. Let $\theta_t|y_{1:t-1} \sim \mathcal{N}(a_t, R_t)$, using Equation (42) and applying Lemna (1), a_t and R_t are given by:

$$\begin{aligned} a_t &= \mathbf{E}(\theta_t|y_{1:t-1}) = \mathbf{E}\Big[\mathbf{E}(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}\Big] \\ &= \mathbf{E}(G_t\theta_{t-1}|y_{1:t-1}) = G_t m_{t-1} \end{aligned}$$

and:

$$\begin{aligned} R_t &= Var(\theta_t|y_{1:t-1}) \\ &= \mathbf{E}\Big[Var(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}\Big] + Var\Big[\mathbf{E}(\theta_t|\theta_{t-1}, y_{1:t-1})|y_{1:t-1}\Big] \\ &= W_t + G_t C_{t-1} G_t' \end{aligned}$$

2. Let $Y_t|y_{1:t-1} \sim \mathcal{N}(f_t, Q_t)$, using Equation (42), a_t and R_t are given by:

$$\begin{aligned} f_t &= \mathbf{E}(Y_t|y_{1:t-1}) = \mathbf{E}\Big[\mathbf{E}(Y_t|\theta_t, y_{1:t-1}|y_{1:t-1})\Big] \\ &= \mathbf{E}(F_t\theta_t|y_{1:t-1}) = F_t a_t \end{aligned}$$

and:

$$\begin{aligned} Q_t &= Var(Y_t|y_{1:t-1}) \\ &= \mathbf{E}\Big[Var(Y_t|\theta_t, y_{1:t-1}|y_{1:t-1})\Big] + Var\Big[\mathbf{E}\left(Y_t|\theta_t, y_{1:t-1}\right)|y_{1:t-1}\Big] \\ &= V_t + F_t R_t F_t' \end{aligned}$$

3. In order to compute the filtering distribution at time t, we have to apply the Bayes formula (Lemna (2)) to combine the prior $p(\theta_t|y_{1:t-1})$ and the likelihood $p(y_t|\theta_t)$ as follows:

$$\begin{aligned} p(\theta_t|y_{1:t}) &= p(\theta_t|y_{1:t-1}, y_t) = \frac{p(y_t|\theta_t, y_{1:t-1}) \cdot p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \\ &= \frac{p(y_t|\theta_t) \cdot p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \\ &\propto p(y_t|\theta_t) \cdot p(\theta_t|y_{1:t-1}) \end{aligned}$$

The posterior density is then proportional to the product of the likelihood $p(y_t|\theta_t)$ and the prior density $p(\theta_t|y_{1:t-1})$.

As previously mentioned, in the DLM case, all the distributions are Gaussian. More specifically, $Y_t|\theta_t \sim \mathcal{N}\left(F_t\theta_t, V_t\right)$ and $\theta_t|y_{1:t-1} \sim \mathcal{N}\left(a_t, R_t\right)$, which leads to:

$$\begin{aligned} p(y_t|\theta_t) &= (2\pi)^{-\frac{1}{2}}\left|V_t\right|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(Y_t - F_t\theta_t)'V_t^{-1}(Y_t - F_t\theta_t)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\Big(Y_t'V_t^{-1}Y_t - 2\theta_t'F_t'V_t^{-1}Y_t + \theta_t'F_t'V_t^{-1}F_t\theta_t\Big)\right\} \end{aligned}$$

and:

$$\begin{aligned} p(\theta_t|y_{1:t-1} &= (2\pi)^{-\frac{1}{2}}\left|R_t\right|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(\theta_t - a_t)'R_t^{-1}(\theta_t - a_t)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\Big(\theta_t'R_t^{-1}\theta_t - 2\theta_t'R_t^{-1}a_t\Big)\right\} \end{aligned}$$

It follows:

$$\begin{aligned} p(\theta_t|y_{1:t}) &\propto \exp\left\{-\frac{1}{2}\Big(\theta_t'F_t'V_t^{-1}F_t\theta_t - 2\theta_t'F_t'V_t^{-1}Y_t\Big)\right\} \cdot \exp\left\{-\frac{1}{2}\Big(\theta_t'R_t^{-1}\theta_t - 2\theta_t'R_t^{-1}a_t\Big)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\Big[\theta_t'\left(F_t'V_t^{-1}F_t + R_t^{-1}\right)\theta_t - 2\theta_t'\left(F_t'V_t^{-1}Y_t + R_t^{-1}a_t\right)\Big]\right\} \end{aligned}$$

Thanks to Proposition (1) we recognize here the Kernel of a normal density $\mathcal{N}\left(m_t, C_t\right)$ with $C_t = \left(F_t'V_t^{-1}F_t + R_t^{-1}\right)^{-1}$ and $m_t = C_t\left(F_t'V_t^{-1}Y_t + R_t^{-1}a_t\right)$.

Applying Lemna (3) to this expression with $A = R_t$, $B = F_t'$ and $C = V_t$, leads to:

$$\begin{aligned} C_t &= R_t - R_t F_t' \left(F_t R_t F_t' + V_t \right)^{-1} F_t R_t \\ &= R_t - R_t F_t' Q_t^{-1} F_t R_t \end{aligned}$$

Finally, plugging the final expression of C_t into that of m_t, we get:

$$m_t = a_t + R_t F_t' Q_t^{-1} (Y_t - f_t)$$

■

C.1.3 Dealing with unknown parameters

In the previous section, we presented the basic DLM assuming that F_t, G_t, V_t, and W_t were known. In practice however, this is rarely true. In this paper we have assumed the model matrices to depend on an unknown parameter vector ψ. Two main approaches can then be opposed to estimate ψ, i.e. the frequentist view that relies on the likelihood principle (LP) and the Bayesian approach that considers parameters as random quantities. Let us briefly describe the ML principle and then focus on the practical aspects of the Bayesian standpoint we adopt in this paper.

According to Berger and Wolpert (1984), the LP deals with situations in which Y has a density $p(y)$ (with respect to some measure ν) for all $\psi \in \Psi$. Of crucial importance is the likelihood function for ψ given y, given by $L(\psi) = p(y_1, \ldots, y_T; \psi)$, i.e. the density evaluated at the observed value $Y = y = (y_1, \ldots, y_T)$ and considered as a function of ψ. More specifically, the LP stands that all the information about ψ obtainable from an experiment is contained in the likelihood function for ψ given y. Two likelihood functions for ψ (from the same or different experiments) contain the same information about ψ if they are proportional to one another.

In the case of a DLM, it is convenient to write the joint density of the observations in the form:

$$p(y_1, \ldots, y_T; \psi) = \prod_{t=1}^{T} p(y_t | y_{1:t-1}; \psi)$$

with $p(y_t | y_{1:t-1}; \psi)$ the conditional density of y_t given the data up to time $t - 1$.

We know from above that the terms from the RHS of the equation are Gaussian densities with mean f_t and variance Q_t. We can therefore rewrite the log-likelihood as follows:

$$l(\psi) = \ln \Big[L(\psi) \Big] = -\frac{Tp}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln |Q_t(\psi)| - \frac{1}{2} \sum_{t=1}^{T} (y_t - f_t(\psi))' Q_t(\psi)^{-1} (y_t - f_t(\psi)) \quad (47)$$

The expression can then be numerically maximized to obtain the MLE of ψ as follows:

$$\widehat{\psi} = \underset{\psi}{\operatorname{argmax}}\, l(\psi) \quad (48)$$

Once an MLE is provided for ψ, it can be introduced into the state-space model, assuming that the value will remain constant, and the Kalman filter can be applied, as detailed in the previous section.[37]

37. For more details on the MLE in the context of state space models, see Chapter 7 from Durbin and Koopman (2001).

As recalled by Petris et al. (2009), the common practice of plugging the MLE in the filtering and smoothing recursions suffers from the difficulties in taking properly into account the uncertainty about ψ. The Bayesian approach offers then a more consistent formulation to the problem. Whereas in the previous case, ψ was assumed constant, here, it is regarded as a random quantity, with a known prior density $p(\psi)$. The general hypotheses of state space models for the processes $\{Y_t\}$ and $\{\theta_t\}$ are then assumed to hold *conditionally* on ψ.
This leads to the following condition (to be compared to Equation (39)) for any $n \geq 1$:

$$(Y_1, \ldots, Y_T, \theta_0, \theta_1, \ldots, \theta_T, \psi) \sim p(\theta_0|\psi) \cdot p(\psi) \prod_{t=1}^{T} p(y_t|\theta_t, \psi) \cdot p(\theta_t|\theta_{t-1}, \psi)$$

Proof. We can give the intuition of the proof as follows.
We first express the joint density of Y_t, θ_t, ψ as the product of the conditional density of Y_t, θ_t given ψ with the prior $p(\psi)$:

$$\begin{aligned} p(\theta_{0:T}, y_{1:T}, \psi) &= p(\theta_{0:T}, y_{1:T}|\psi)p(\psi) \\ &= p(y_{1:T}|\theta_{0:T}, \psi) \cdot p(\theta_{0:T}|\psi) \cdot p(\psi) \end{aligned}$$

As $\theta_t|\psi$ is a Markov chain, we can use the result exhibited in Equation (39). We get:

$$p(\theta_{0:T}|\psi) = p(\theta_0|\psi) \prod_{t=1}^{T} p(\theta_t|\theta_{t-1}, \psi)$$

As $Y_t|\psi$ depends only on $\theta_t|\psi$, for all t, we get:

$$p(y_{1:T}|\theta_{0:T}, \psi) = \prod_{t=1}^{T} p(y_t|\theta_t, \psi)$$

Thus, we obtain:

$$p(\theta_{0:T}, y_{1:T}, \psi) = p(\theta_0|\psi) \cdot p(\psi) \prod_{t=1}^{T} p(y_t|\theta_t, \psi) \cdot p(\theta_t|\theta_{t-1}, \psi)$$

■

Given the data $y_{1:t}$, inference on the unknown state θ_s at time s and on the parameters ψ is solved by computing their joint posterior distribution via the recursion *formulae* given in the previous section. We obtain:

$$p(\theta_s, \psi|y_{1:t}) = p(\theta_s|y_{1:t}, \psi) \cdot p(\psi|y_{1:t}) \tag{49}$$

where as previously, one might be interested in $s = t$ in filtering problems, $s \geq t$ for state predictions and $s \leq t$ for smoothing. Here we focus on the filtering case and then assume $s = t$ for what follows.

The marginal conditional density of θ_t is then obtained by averaging Equation (49) with respect to the posterior distribution of ψ given the data $y_{1:t}$:

$$p(\theta_t|y_{1:t}) = \int p(\theta_t|y_{1:t}, \psi) \cdot p(\psi|y_{1:t})d\psi \tag{50}$$

As recalled by Petris et al. (2009), most of the time, one is interested in reconstructing all the unknown state history up to time t and Equation (49) takes the general formulation of:

$$p(\theta_{0:t}, \psi | y_{1:t}) = p(\theta_{0:t} | y_{1:t}, \psi) \cdot p(\psi | y_{1:t}) \tag{51}$$

In principle, the posterior density specified in Equation (51) can be derived from Lemna (2). In practice however, it is most of the time impossible to obtain closed-forms and one has to rely on simulation methods. In particular, importance sampling or Markov Chain Monte Carlo (MCMC) methods provide efficient tools for approximating the posterior distributions of interest.[38] Adopting this second class of approach, the following section gives some details about the procedure we implemented and some estimation results.

C.2 Our model

We present here some proofs and complementary results associated to the Bayesian analysis presented in Section 2.2.2.

C.2.1 Conjugate priors

In Section 2.2.2 we define our prior densities as follows:

Autoregressive parameters ψ_1:

$$\begin{aligned} (\phi_1, \phi_2) &\sim \mathcal{N}\left(0, \left(\frac{2}{3}\right)^2\right) \times \mathcal{N}\left(0, \left(\frac{1}{3}\right)^2\right), \text{restricted to } \Omega \text{ with:} \\ \Omega &= \{\phi_1 + \phi_2 < 1, \phi_1 - \phi_2 > -1, |\phi_2| < 1\} \end{aligned}$$

Precision parameters ψ_2:

$$\sigma_\varepsilon^{-2}, \sigma_z^{-2}, \sigma_u^{-2} \sim \mathcal{G}\left(\frac{a^2}{b}, \frac{a}{b}\right) \text{ with } \mathbf{E}\left(\sigma_i^{-2}\right) = a \text{ and } Var\left(\sigma_i^{-2}\right) = b, \forall i \in \{\varepsilon, z, u\}$$

After calculations, we obtain the following conditionally independent and gamma-distributed precisions:

$$\sigma_\varepsilon^{-2} | \cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,1} - (G_t\theta_{t-1,1})_1\right)^2\right) \tag{52}$$

$$\sigma_z^{-2} | \cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,2} - (G_t\theta_{t-1,2})_2\right)^2\right) \tag{53}$$

$$\sigma_u^{-2} | \cdots \sim \mathcal{G}\left(\frac{a^2}{b} + \frac{T}{2}, \frac{a}{b} + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,3} - (G_t\theta_{t-1,3})_3\right)^2\right) \tag{54}$$

with: $\sigma_i^{-2} | \sigma_j^{-2}, \sigma_k^{-2}, \theta_{0:t}, y_{1:t} \sim \sigma_i^{-2} | \ldots, \forall (i,j,k) \in \{\varepsilon, z, u\}, i \neq j, i \neq k,$ and $j \neq k$

38. For more insights on the Bayesian approach to DLM, see Chapters 8, and concerning importance sampling, Chapters 11 and 13 from Durbin and Koopman (2001). For a broader description of MCMC methods applied to the DLM case, see Chapters 4 and 5 in Petris et al. (2009).

Let us give some elements to prove the last result.

Proposition 2 Kernel of a Gamma distribution
Assuming that $Y \sim \mathcal{G}(\alpha, \beta)$, the density of Y, $p(y)$, is given by:

$$\begin{aligned} p(y) &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot y^{\alpha-1} \cdot e^{-\beta y} \\ &\propto y^{\alpha-1} \cdot e^{-\beta y} \end{aligned}$$

Proof. Let us first rename the vector of states θ_t that, in our case, is of dimension 4 as follows: $\theta_t = (\theta_{t,1}, \theta_{t,2}, \theta_{t,3}, \theta_{t,4})'$. According to the description in Section 2.2.1, this leads to $\theta_{t,1} = s_t^P$, $\theta_{t,2} = \delta_t$, and $\theta_{t,3} = s_t^C$. In addition, as it is more convenient to work with precisions rather than variances, we associate to those four components the following $\phi_i = \sigma_i^{-2}$, $\forall i \in \{1,2,3,4\}$, which correspond to σ_ϵ^{-2},σ_z^{-2},σ_u^{-2}, and 0 respectively.
For all $i \in \{1,2,3\}$, the conditional density of $\phi_i | \ldots$ can be expressed as follows:

$$p(\phi_i|\ldots) \propto p(\theta_{0:T}|y_{1:T}, \phi_i, \phi_j, \phi_k) \cdot p(\phi_i|y_{1:T}, \phi_j, \phi_k) \tag{55}$$

$$\propto \prod_{t=1}^{T} p(\theta_{i,t}|\theta_{i,t-1}, \phi_i) \cdot p(\phi_i) \tag{56}$$

We know that:

$$\begin{aligned} \theta_{t,i}|\theta_{t-1,i}, \phi_i &\sim \mathcal{N}\left(G_t \theta_{t,i}, \phi_i^{-1}\right) \\ \phi_i &\sim \mathcal{G}(\alpha_0, \beta_0) \text{ with } \alpha_0 = \frac{a^2}{b}, \beta_0 = \frac{a}{b} \end{aligned}$$

Thus, by application of Propositions 1 and 2, we obtain:

$$\begin{aligned} \prod_{t=1}^{T} p(\theta_{t,i}|\theta_{t-1,i}, \phi_i) &\propto \phi_i^{\frac{T}{2}} \cdot \exp\left\{-\frac{1}{2}\phi_i \sum_{t=1}^{T} \left(\theta_{t,i} - (G_t\theta_{t-1,i})_i\right)^2\right\} \\ p(\phi_i) &\propto \phi_i^{\alpha_0 - 1} \cdot \exp\{-\beta_0 \phi_i\} \end{aligned}$$

Finally, it comes:

$$p(\phi_i|\ldots) \propto \phi_i^{\frac{T}{2}+\alpha_0-1} \cdot \exp\left\{-\phi_i\Big(\frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,i} - (G_t\theta_{t-1,i})_i\right)^2 + \beta_0\Big)\right\}$$

Thanks to Proposition 2, we recognize the Kernel of a Gamma distribution of parameters $\alpha = \alpha_0 + \frac{T}{2}$ and $\beta = \beta_0 + \frac{1}{2}\sum_{t=1}^{T}\left(\theta_{t,i} - (G_t\theta_{t-1,i})_i\right)^2$. ■

C.2.2 Hybrid sampler

As previously mentioned, it frequently happens that posterior distributions of the unknown parameters (ψ) turn out to be analytically intractable. In such a context, it is impossible to derive the marginal distribution of a particular parameter or subset of parameters (ψ_1, ψ_2). Consequently, one needs to implement sampling methods. Two broad classes of methods exist in the literature: (i) importance sampling approaches and (ii) Markov chain Monte Carlo (MCMC) methods, both of which relying on the same basic principle that consists in approximating an intractable integral or sum using samples from some distribution. If the distribution under scrutiny is a closed (known) distribution from the one (unknown) targeted, we deal with the importance sampling approach which consists then in drawing samples from the closed density with appropriate weights (see pp. 76-83 from Bauwens et al. (1999) or pp. 110-119 from Geweke (2005) for more details). Conversely, MCMC methods rely on the derivation of an ergodic Markov chain whose limit distribution is exactly the one targeted, (see pp. 83-93 from Bauwens et al. (1999) or pp. 119-127 from Geweke (2005) for more insights). The literature is generally divided between those who call for the use of importance sampling, arguing that this method is more transparent and computationally more convenient than MCMC, [39] and those who defend the great liberty of MCMC approaches, pointing that it is sometimes more difficult to find a closed distribution with the associated "good" properties, [40] than sampling univariate conditional densities of the targeted one. Moreover, as Bauwens et al. (1999) recalls, whereas importance sampling is problem specific (the choice of a good approximation for the posterior density depends on the model and sometimes also the sample), MCMC methods can be applied to a wide range of problems indifferently, making them somehow easier to implement and less intensive in research time. As Persi Diaconis notice, [41] there has been very little rigorous work comparing these two algorithms and thus no consensus. In this paper, dealing with a complex problem including the search for stationarity conditions for our cyclical component $\theta_{t,3} = s_t^C$, we chose to rely on the second approach to derive the posterior distributions of interest and the relative statistical summaries (means, variances). Following Petris et al. (2009), we implement an hybrid sampler to derive the conditional density of ψ_1, the vector containing the autoregressive parameters, $\theta_{0:n}$, the states (including the medium-run equilibrium value $\theta_{t,1} = s_t^P$), and ψ_2 which contains the precision (variance) parameters. More specifically, to those three groups of parameters are associated three different sampling methods, namely, an Adaptive rejection Metropolis sampling (ARMS), a Forward Filtering Backward Sampling (FFBS), and finally a standard Gibbs sampling. We briefly describe each of those approaches in the following lines. [42]

Gibbs sampling The Gibbs sampler is required to be able to simulate the "full" conditional densities of $p(\psi)|y_{1:T}, \theta_{0:T}$. This means that one must be able to partition ψ into k blocks such that $\forall i \in \{1, \ldots, k\}$, $p(\psi_i|\psi_{-i})$ can be directly simulated.

One first starts from an arbitrary point $\psi^{(0)} = (\psi_1^{(0)}, \ldots, \psi_k^{(0)})$ in the parameter space and "updates" one component at a time by drawing ψ_i, $i \in \{1, \ldots, k\}$ from the relevant conditional distribution, according to the scheme described in Box 9.

Forward Filtering Backward Sampling As Petris et al. (2009) recall, in a Gibbs sampling from $p\big(\theta_{0:T}, \psi|y_{1:T}\big)$,

39. See for instance Durbin and Koopman (2001) on that point at the end of Chapter 8.

40. See Geweke (1989).

41. A film of his keynote lecture on "Importance Sampling vs. Markov chain Monte Carlo" given for the 'Markov Chains in Algorithms and Statistical Physics' workshop in the Mathematical Sciences Research Institute from the University of Berkeley, on January 31, 2005 is available at: `http://www.msri.org/communications/vmath/VMathVideos/VideoInfo/1859/show_video`.

42. For more details, see Petris et al. (2009) or Gamerman and Freitas Lopes (2006).

Figure 9: The Gibbs algorithm

1. Initialize the starting point: $\psi^{(0)} = (\psi_1^{(0)}, \ldots, \psi_k^{(0)})$;

2. For $j = 1, \ldots, N$:

   2.1. generate $\psi_1^{(j)}$ from $p\big(\psi_2^{(j-1)}, \ldots, \psi_k^{(j-1)} \big| y_{1:T}, \theta_{0:T}\big)$

   2.2. generate $\psi_2^{(j)}$ from $p\big(\psi_1^{(j)}, \psi_3^{(j-1)}, \ldots, \psi_k^{(j-1)} \big| y_{1:T}, \theta_{0:T}\big)$

   ⋮

   2.k. generate $\psi_k^{(j)}$ from $p\big(\psi_1^{(j)}, \ldots, \psi_{k-1}^{(j)} \big| y_{1:T}, \theta_{0:T}\big)$

one needs to simulate from the full conditional densities $p\big(\psi|y_{1:T}, \theta_{0:T},\big)$ and $p\big(\theta_{0:T}|y_{1:T}, \psi,\big)$. While the first density is problem specific (and requires either a Gibbs or a Metropolis Hastings step, as described in the following paragraph), the general expression of the latter can be derived from an efficient algorithm called the Forward Filtering Backward Sampling (FFBS), primarily developed by Frühwirth-Schnatter (1994) among others. The idea is the following:

Figure 10: The forward filtering backward algorithm

1. We can write the joint distribution of $\theta_{0:T}|y_{1:T}, \psi$ as follows:

$$\begin{aligned} p\big(\theta_{0:T}|y_{1:T}, \psi\big) &= p\big(\theta_0, \theta_{1:T}|y_{1:T}, \psi\big) \\ &= p\big(\theta_0|\theta_{1:T}, y_{1:T}, \psi\big) \cdot p\big(\theta_{1:T}|y_{1:T}, \psi\big) \\ &= p\big(\theta_0|\theta_{1:T}, y_{1:T}, \psi\big) \ldots p\big(\theta_{T-1}|\theta_T, y_{1:T}, \psi\big) \cdot p\big(\theta_T|y_{1:T}, \psi\big) \end{aligned}$$

   The last factor is $p\left(\theta_T|y_{1:T}, \psi\right)$ the filtering distribution as described *infra*, which follows a normal distribution $\mathcal{N}\left(m_T, C_t\right)$.

2. In order to obtain a draw from the distribution on the LHS, one can start by drawing θ_T from a $\mathcal{N}\left(m_T, C_T\right)$ and then, for $t = T-1, T-2, \ldots, 0$ recursively draw θ_t from $p\big(\theta_t|\theta_{t+1:T}, y_{1:T}, \psi\big)$. It can be shown that this last expression follows a normal distribution characterized by h_t and H_t, with $h_t = m_t + C_t G'_{t+1} R_{t+1}^{-1}(\theta_{t+1} - a_{t+1})$ and $H_t = C_t - C_t G'_{t+1} R_{t+1}^{-1} G_{t+1} C_t$. [a]

a. This last expression derives from the smoothing condition. For more details, see Petris et al. (2009) and Frühwirth-Schnatter (1994).

Adaptive rejection Metropolis Sampling As introduced by Walter (Wally) Gilks, the "father" of this sampling approach, Adaptive rejection Metropolis sampling (ARMS) is a method for efficiently sampling from complicated univariate densities, such as typically occur in applications of Gibbs sampling (see Gilks et al., 1995). More specifically, as described by Petris et al. (2009) rejection sampling is an algorithm that allows one to generate a random variable from a target distribution p

by drawing from a different proposal distribution f and then accepting with a specific probability. It consists in the following steps:

Figure 11: The ARM algorithm

1. Assume there is a constant C such that $p(\psi) \leq Cf(\psi)$ for every ψ and define $r(\psi) = \frac{p(\psi)}{Cf(\psi)}$ so that $0 \leq r(\psi) \leq 1$;
2. Generate two independent random variables:
   (a) $U \sim \mathcal{U}_{[0,1]}$
   (b) $V \sim f$
3. If $U \leq r(V)$ set $\psi = V$, otherwise repeat the process

This algorithm allows then to draw V from $f(\psi)$ and accept V as a fraw from $p(\psi)$ with probability $r(V)$. It can be shown that if the support of p is included in that of f, then the algorithm terminates in a finite time. For more details, see: http://www.amsta.leeds.ac.uk/~wally.gilks/adaptive.rejection/web_page/Welcome.html

C.2.3 Estimations

MLE estimates of ψ Here are the estimates for the parameters for both cases under scrutiny (convergence vs. divergence *scenarii*):

Table 2: Maximum likelihood estimates

	ψ_1		ψ_2		
Model	ϕ_1	ϕ_2	σ^2_ϵ	σ^2_z	σ^2_u
Convergence	1.781	−0.798	$1.862.10^{-3}$	$1.295.10^{-4}$	$9.386.10^{-3}$
Divergence	1.799	−0.799	$4.110.10^{-4}$	$4.066.10^{-4}$	$1.027.10^{-2}$

Bayesian estimates of ψ Here we exhibit the Bayesian estimates for $\psi_1 = (\phi_1, \phi_2)$ and $\psi_2^{-1} = (\sigma^2_\epsilon, \sigma^2_z, \sigma^2_u)$ and the associated standard errors in parentheses, depending on the number of iterations for the sampler ($MCMC$), with a burn-in $B = 100$ in each case.

Initial values for ψ We assumes either MLE estimates (scenario 1, when agents are right) or parameters that are assumed to be far from the MLE estimates, more specifically, we assume $\phi_i = 0.9$ and $\ln(\sigma^2_j) = -10 \Leftrightarrow \sigma^2_j = 4.534.10^{-5}$, $\forall i \in \{1, 2\}$ and $j \in \{\epsilon, z, u\}$

Table 3: Bayesian estimates

	ψ_1		ψ_2^{-1}		
MCMC	ϕ_1	ϕ_2	σ^2_ϵ	σ^2_z	σ^2_u
1000	0.734	-0.205	$1.28.10^{-3}$	$5.22.10^{-3}$	$1.10.10^{-3}$
	(0.010)	(0.007)	$(4.66.10^{-5})$	$(6.84.10^{-5})$	$(3.84)^{-5}$
6000	0.747	-0.219	$1.21.10^{-3}$	$5.21.10^{-5}$	$1.13.10^{-3}$
	(0.004)	(0.003)	$(1.67.10^{-5})$	$(2.78.10^{-5})$	$(1.60.10^{-5})$
10000	0.753	-0217	$1.21.10^{-3}$	$5.15.10^{-5}$	$1.17.10^{-5}$
	(0.003)	(0.002)	$(1.31.10^{-5})$	$(2.23.10^{-5})$	$(1.30.10^{-5})$