Universite catholique de Louvain
Faculté des Sciences Économiques, Sociales, Politiques et de Communication
Departement des Sciences Économiques
Institut de Recherches Économiques et Sociales

Four Essays on Uncertainty and Expectations in Macroeconomics

Pablo Aguilar

These à présenter en vue de l'obtention du grade de docteur

en sciences economiques et de gestion

Composition du Jury:

Luca Pensieroso (Promoteur), UCLouvain

Francesca Monti (Président), UCLouvain

Yuliya Rychalovska, UNamur

Jesús Vázquez, UPV/EHU

Raf Wouters, NBB

Louvain-la-Neuve, Belgique, September 2022

Acknowledgments

I would like to take this opportunity to sincerely thank my thesis advisor, Luca Pensieroso, for being there all these (many) years. From you I had always lots of intuition, care for the details, and economic, model and data discipline. I also want to thank Raf Wouters for his guidance all along, many fellows have benefited from your comments and feedbacks, I'm happy to be one of them. I would like to extend my gratitude to the rest of the jury as well.

This has been a long journey and I would like to share my gratitude to all that have shared it with me through the different stages: my colleagues at Louvain, at Frankfurt and now in Madrid.

Finally I want to express my gratitude to Mercedes, my wife, for accompanying me all these years and to my family, that has always encourage me pursue this PhD. I couldn't be happier about it!!

Contents

General introduction

Traditional economic analysis, in the form of rational expectation models, assumes that agents have perfect information about the state of the economy and are capable of understanding the nature of different shocks, their duration and transmission to the rest of variables in the economy. However, in reality, agents face economic decisions with limited information and under have great uncertainty about the economic outlook. Bearing this in mind, this thesis addresses the importance of uncertainty and expectations in the economy by providing four pieces of analysis.[1] The first two chapters explore how agents use financial information in the presence of limited information when forming expectations and how financial markets may capture and transmit economic uncertainty. The third chapter provides a textual-analysis based indicator to measure the real-time impact of uncertainty on economic activity during periods with exceptionally high uncertainty such as the COVID-19 crisis, while the fourth chapter explores the role of uncertainty through expectations as a driving factor during large crisis.

More into detail, the first chapter analyzes the importance of financial information when agents have to forecast macroeconomic outcomes and how this process can feed back into the economy through the lens of a Dynamic Stochastic General Equilibrium (DSGE) model with agents under adaptive learning expectations. Adding financial information allows the modelling of the term structure of interest rate under the Expectation Hypothesis of the term structure and the incorporation of changes in the short-term spread into agents' forecasting models. This leads to financial information becoming an additional source of business cycle persistence through the expectation channel and amplifies the impact of term premium innovations in the economy, normally minor in this sort of models. In addition, the estimated

[1] The first chapter, *An estimated DSGE model with Learning based on term structure information* , with Jesús Vázquez (UPV/EHU) has been published in Macroeconomic Dynamics (2021), Vol. 25, N.10. and likewise for the third chapter *Can news help measure economic sentiment? An application in COVID-19 times,* With Corinna Ghirelli, Matías Pacce and Alberto Urtasun (BdE), published in Economics Letters (2021), Vol. 199.

model for the United States (U.S.) shows a similar fit to the postwar period than rational expectations models with a lower degree of persistence needed in the endogenous variables and increases the capacity of the model in reproducing the hump-shaped pattern of U.S. inflation over the last 50 years.

The second chapter deepens in the importance of financial information in DSGE models with adaptive learning expectations by introducing the term structure of interest rates through bond yields that are priced with separate Euler equations. This approach enables the estimation of a measure of the term premium associated with the 10-year US Treasury bond. Considering the standard consumption-based asset pricing equation associated with each maturity leaves the Expectation Hypothesis of the term structure from the previous chapter and explicitly turns into multiple-period-ahead expectations. This enhances the expectations channel in the model as pricing with the Euler equation relates the term premium to a wedge between the optimal yield and the expectations about future consumption and inflation. In addition, it permits defining simple, and forward-looking forecasting rules that only depends on term structure information. The estimation results show that the extended model with term structure and simple forecasting rules provides a much better fit to the data than the rational expectations version, and a measure of the term premium that is highly correlated with other term premiums from the relevant literature. As an extension exercise the model also includes survey data on inflation and consumption expectations to verify the similar statistical properties of model implied expectations with the data.

The third chapter presents a methodology to measure the real-time impact of uncertainty in the economy during crisis. Based on textual analysis this chapter constructs an economic sentiment indicator capturing the economic tone of news published in the Spanish press. A major advantage of constructing a newspaper-based indexes compared to traditional (monthly) confidence indicators is that each day it incorporates new information in the index. This becomes crucial when important shocks occur in the middle of the month or far from the date at which the monthly confidence indicators are released. Using nowcast techniques this chapter shows how this daily indicator becomes relevant in tracking economic activity in real-time and predicting the COVID-19 recession from an earlier date.

Finally, the fourth chapter shows how uncertainty affects agents' expectations during large crisis and specially during the recovery phase. This chapter introduces adaptive learning in a stylized Dynamic General Equilibrium model of the U.S. Great Depression of the 1930s, thereby contributing to assessing the quantitative importance of uncertainty as a driving factor of the Depression. Results from simulations show that the calibrated model with adaptive learning outperforms its rational-expectations counterpart, especially for what concerns the explanation of the slow recovery from the Depression, pointing to uncertainty as an important factor delaying the recovery, that is not capture under rational expectations models. Considering uncertainty in the form of a light departure from rational expectations is a promising way to take its important role in big crises into account, without introducing animal spirits or forsaking the analytical advantage and quantitative strength of DGE modelling. These results are robust to different specifications of the learning dynamics, while the data mimicking ability increases as the model gets further away from rational expectations.

Part I

An estimated DSGE model with Learning based on term structure information

1 Introduction

Agents can learn from financial markets to forecast macroeconomic outcomes. At the same time, learning dynamics can feed back into both the macroeconomy and financial markets. This paper introduces the term structure of interest rates into an estimated medium-scale dynamic stochastic general equilibrium (DSGE) model with adaptive learning (AL) expectations. Its aim is twofold. First, we analyze the role of the term structure of interest rates in the learning process of economic agents and show that this feature may take over other explanations of aggregate persistence. Second, we study how term structure innovations are transmitted into the macroeconomy under AL.

We build on the medium-scale AL model of Slobodyan and Wouters (2012a) based on small forecasting models by allowing agents explicitly to use data contained in the term structure of interest rates to fully characterize the formation of their expectations. This results in a major deviation from most AL models. Estimated AL models typically use *strongly revised* data, whereas actual learning dynamics must be driven by data available to agents when forming their expectations in real time.[1] Our approach overcomes this shortcoming because

With Jesús Vázquez (UPV/EHU), published in Macroeconomic Dynamics (2021), Vol. 25, N.10.

[1]There are a few exceptions that address this important shortcoming in the AL literature. Milani (2011) focuses on real-time data on output and inflation and forecasts of them from the Survey of Professional Forecasters (SPF) recorded in real time when estimating a small-scale DSGE model. However, he ignores revised data on macroeconomic variables, which more accurately describe the actual economy, when estimating and assessing the model's fit. Orphanides and Wei (2012) also use real-time data, but focus on a VAR model

term structure information provides real-time information, which is not revised, about the behavior of the aggregate economy.[2] Apart from this distinctive feature, it is important to remark that our AL approach still shares an important feature with other AL approaches: The process of updating AL coefficients depends on actual values of forward-looking variables (e.g. aggregate consumption and inflation) because they are updated based on the forecast errors of these observable variables.

From a theoretical perspective, the term structure of interest rates can be viewed as summarizing information about expectations on both the macroeconomy and monetary policy. Thus, by introducing the term structure of interest rates, we consider a new channel where multi-period-ahead forecasts matter for household behavior. This feature is in line with ideas in Marcet and Sargent (1989), Preston (2005), Eusepi and Preston (2011), and Sinha (2016) of modeling optimal decisions conditional on multi-period forecasts. The rationale behind our approach is further supported by a large body of empirical literature (Fama, 1990; Mishkin, 1991; Estrella and Mishkin, 1997; Ang, Piazzesi and Wei, 2006) that evidences the ability of the term spread to predict both inflation and economic activity.

The small forecasting models assumed in this paper are very simple, but they perform well because the time-varying intercepts in them capture the long-run features of data well (e.g. the switching inflation trend in the early 1980's) whereas the time-varying coefficients associated with term spreads capture short-run movements relatively well (e.g. the growth rates in consumption and investment).

Our estimation results show that, compared with the rational expectations (RE) version of the model, including the term structure in the AL model provides a similar fit to post-war US

rather than a DSGE model. Similarly, Slobodyan and Wouters (2013) use both real-time data and survey expectations data in a univariate model of inflation.

[2]The idea of using only term structure information to predict business cycle conditions is not new (McCallum, 1994). The use of only term structure information to characterize agent's expectations can certainly be seen as somewhat restrictive because agents indeed observe additional relevant information. Below we show that considering additional (revised) data does not improve model fit once term structure information is included in small forecasting models. Moreover, in a follow-up paper Aguilar and Vázquez (2018) show that adding real-time data into small forecasting models does not improve the overall performance of the DSGE model estimated.

data but results in a much lower degree of endogenous persistence associated with price and wage stickiness and the elasticity of the cost of adjusting capital. Our findings corroborate the fall in the elasticity of the cost of adjusting capital and price stickiness found in Slobodyan and Wouters (2012a) under AL.[3] The rationale behind these findings is that including the term structure in the DSGE model brings multi-period forecasting into the representative agent decision problem, which results in a much stronger persistence driven by learning dynamics than that induced by the one-period-ahead forecasts considered in their paper. Moreover, our findings are in line with those of Eusepi and Preston (2011) in a prototype real business cycle model, where multi-period forecasting results in much more persistent aggregate dynamics. Our empirical analysis also shows that including the term structure in the AL model provides additional support for important features found by Slobodyan and Wouters (2012a) in their estimated medium-scale DSGE model with AL. In particular, the AL model with term structure also reproduces the hump-shaped pattern of U.S. inflation over the last fifty years. Furthermore, the AL model estimated shows that term spread innovations are a major source of fluctuations in nominal variables. This finding stands in sharp contrast to the lack of transmission of term spread shocks to the macroeconomy under RE.

The rest of the paper is organized as follows. Section 2 connects the contribution of this paper to the related literature. Section 3 introduces the term structure of interest rates into the medium-scale AL model. Section 4 discusses the main estimation results. Section 5 studies the transmission of shocks with a focus on term premium shocks. Section 6 assesses the robustness of estimation results across alternative formulations of the forecasting models. Section 7 concludes.

[3] Milani (2007) reached qualitatively similar conclusions, but he studied a small-scale DSGE model, where many sources of aggregate persistence were neglected.

2 Related literature

Following the learning approach of Evans and Honkapohja (2001), this paper also builds on the growing literature that investigates the role of AL as an alternative to the assumption of RE in the analysis of DSGE models.[4] Recent papers (Orphanides and Williams, 2005a; Milani, 2007, 2008, 2011; and Eusepi and Preston 2011) focus on small-scale DSGE models whereas Slobodyan and Wouters (2012a,b) introduce AL into the medium-scale DSGE model of Smets and Wouters (2007), hereinafter referred as the SW model.[5] The first group of papers considers that agents' expectations are based on a linear function of the state variables of the model whose learning coefficients are updated in each period under a gain rule (i.e. the minimum state variable approach), but Slobodyan and Wouters (2012a,b) consider an AL model with agents who form expectations using small forecasting models. Small forecasting models typically assume that agents form their expectations based on the information provided by observable endogenous variables (e.g. those showing up in Euler equations). Considering small forecasting models based only on observable variables is arguably a more appealing approach to AL than the minimum state variable approach since the latter requires agents to know the true model economy (e.g. they know what the state variables are and they perfectly observe them).[6]

Our paper is also related to others that study the interaction between learning dynamics and term structure information (Dewachter, Iania and Lyrio, 2011; Sinha, 2015, 2016). Those

[4]This paper is also related to another strand of the literature (Hördahl, Tristani and Vestin, 2006; Rudebusch and and Wu, 2008; Bekaert, Cho and Moreno, 2010) that seeks to link small-scale new Keynesian monetary model dynamics with the term structure of interest rates under RE. De Graeve, Emiris and Wouters (2009) show evidence of the importance of considering medium-scale RE-DSGE models in order to understand the links between the term structure and the aggregate economy.

[5]There is also a large body of macroeconomic literature that analyzes deviations from the RE assumption in the context of small-scale models. This literature includes, among others, the rational inattention approach (Sims, 2003; Adam, 2007; Mackowiach and Wiederholt, 2009), the sticky information approach (Mankiw and Reis, 2002; Reis, 2009), and several approaches that deal with imperfect information issues (Svensson and Woodford, 2004; Coenen, Levin and Wieland, 2005; Levine, Pearlman, Perendia and Yang, 2012; Pruitt, 2012; Casares and Vázquez, 2016).

[6]Other papers (Adam, 2005; Orphanides and Williams, 2005b; Branch and Evans, 2006; Eusepi and Preston 2011; Hommes and Zhu, 2014; Ormeño and Molnár, 2015) have also provided support for the use of small forecasting models on several grounds such as their relative forecast performance and their ability to approximate the SPF well.

papers mainly focus on how learning dynamics shape the yield curve, but our paper focuses on how term structure information influences learning dynamics to induce a strong source of aggregate persistence.

Deviating from the RE assumption by considering AL based on small forecasting models is appealing for two main reasons. First, in reality agents face incomplete knowledge about the economy, which is at odds with the full information approach assumed under RE. Moreover, gathering and processing information is costly, so economic agents are likely to rely on a small set of variables when trying to figure out the relevant economic environment in their decision processes. Second, AL typically features a sluggish reaction to exogenous and latent shocks that hit the economy, which provides an alternative competing hypothesis on aggregate persistence (Milani, 2007, and Slobodyan and Wouters, 2012a).

Unfortunately, any form of deviation from the RE assumption studied in the literature is arguably arbitrary, and therefore requires further assessment. This is not a simple task. As suggested by Adam and Marcet (2011) and Slobodyan and Wouters (2012a), considering actual data on private sector expectations available through surveys or forward-looking variables, such as asset prices, might be very useful in disciplining expectation formation.[7] In this paper we focus on term structure information rather than on the SPF information used in the related literature because the former provides a much broader (market) view of agents' expectations. Moreover, we assume that the expectations hypothesis of the term structure of interest rates holds under AL and focus on the short end of the term structure. These two features can be viewed as somewhat restrictive, but we introduce them as a way of isolating the contribution of term structure information in characterizing AL. Thus, by imposing the expectation hypothesis of the term structure we downplay the deviation of AL from RE because this hypothesis holds for the (linearized version of the) SW model under RE. Similarly, by focusing on the short end of the yield curve we minimize the deviation of

[7] In this vein, recent papers that introduce AL into DSGE models typically use the SPF to include additional observables in order to discipline agents' beliefs (e.g. Milani, 2011, and Ormeño and Molnár, 2015) or to assess the performance of AL expectations as in Eusepi and Preston (2011) and Slobodyan and Wouters (2012a).

our AL model, which includes a four-quarter-ahead expectation horizon, from the AL model of Slobodyan and Wouters (2012a,b), which considers only one-quarter-ahead expectations. In a follow-up paper, Aguilar and Vázquez (2018), we relax these two assumptions, which enables us to further assess the relative importance of term structure information on AL by considering the whole yield curve.[8]

3 An adaptive learning model with term structure

This paper investigates the contribution of the term structure to the characterization of the agents' learning process. Our model builds on the SW model under AL suggested by Slobodyan and Wouters (2012a) in two important directions. First, we extend the model to account for the term structure of interest rates. Second, we assume that the information in the term structure, which is observed in real time, is the only information that enters the small forecasting models of all forward-looking variables (i.e. those involving expectations in the estimated DSGE model). This standard medium-scale DSGE model contains both nominal and real frictions that affect the choices of households and firms, and we briefly present this model below. However, our main focus is on the extensions related to both the term structure and AL. The complete log-linearized version of the SW model extended with AL and the term structure is presented in the Appendix together with a table describing parameter notation.

[8] Furthermore, Aguilar and Vázquez (2018) consider real-time inflation as well as SPF data on consumption growth and inflation forecasts to discipline agents' expectations. Nonetheless, the use of consumption growth forecasts from the SPF force us to restrict our attention to the Great Moderation period (1984-2007) in that paper because this time series starts in 1981:3. The sample period is much shorter than the one considered (1966-2007) in this paper, which is similar to that studied in Slobodyan and Wouters (2012a,b) for the purposes of comparison. Other important differences between these two papers that show alternative AL modeling choices are highlighted below.

3.1 The SW model

Households maximize their utility, which depends on their levels of consumption relative to an external habit component and leisure. Labor supplied by households is differentiated by a union with monopoly power that sets sticky nominal wages à la Calvo (1983). Households rent capital to firms and decide how much capital to accumulate depending on the capital adjustment costs that they face. Intermediate firms use capital and differentiated labor to produce differentiated goods and set prices à la Calvo. In addition, both prices and wages are partially indexed to lagged inflation when they are not re-optimized, introducing another source of nominal rigidity. As a result, current prices depend on current and expected marginal costs and past inflation whereas current wages are determined by past and expected future inflation and wages. Following Slobodyan and Wouters (2012a), we assume a Taylor-type rule where the short-term nominal interest rate reacts to inflation and to both the level and the growth rate of the output gap, where the latter is defined as the level of output relative to the underlying productivity process, rather than the natural output level used in the SW model.[9] In addition, we assume that the interest rate policy reacts to term structure information, which is in line with agents' learning processes. Finally, the model contains eight stochastic disturbances associated with technology shocks, price and wage markup shocks, and demand-side shocks, including policy interest rate and term premium shocks.

3.2 The term structure extension

This section introduces the term structure into the SW model. Following De Graeve, Emiris and Wouters (2009) and Vázquez, María-Dolores and Londoño (2013), we extend the DSGE model by explicitly considering the interest rates associated with alternative bond maturities indexed by j (i.e. $j = 1, 2, ..., n$). From the first-order conditions characterizing the optimal decisions of the representative consumer, we obtain the standard consumption-based asset

[9]This assumption avoids the modeling of the flexible economy, which includes many additional forward-looking variables. In the rest of the paper we continue to refer to the SW model even though we consider an alternative measure of the output gap.

pricing equation associated with each maturity:

$$E_t\left[\beta^j \frac{U_C(C_{t+j},L_{t+j})\left(exp(\varepsilon_t^{\{j\}})(1+R_t^{\{j\}})\right)^j}{U_C(C_t,L_t)\prod_{k=1}^{j}(1+\pi_{t+k})}\right]=1,\ for\, j=\ 1,\ 2,\ ...,\ n,$$

where E_t stands for the RE or the AL operator depending on the scenarios analyzed below, β is the discount factor, U_C denotes the marginal utility consumption, C_t, L_t, π_t, $R_t^{\{j\}}$, and $\varepsilon_t^{\{j\}}$ denote consumption, labor, the rate of inflation, the nominal yield and the risk premium shock associated with a j-period maturity bond, respectively. The inclusion of a risk premium shock for each maturity is in line with the view of many authors of interpreting the gap between the pure-expectations-hypothesis-implied yield, $R_t^{\{j\}}$, and the observed yield as a measure of fluctuations in the term premium (e.g. De Graeve, Emiris and Wouters, 2009).[10] Moreover, since we focus on US Treasury bond yields in the empirical analysis carried out below, term premium shocks can be viewed as a convenience yield term (Krishnamurthy and Vissing-Jorgensen, 2012; Greenwood et al., 2015; Del Negro et al., 2017) defined as a risk premium associated with the safety and liquidity features of treasury bonds relative to assets with the same payoff and maturity, but without such outstanding properties. Furthermore, the introduction of term premium shocks into the model captures the imperfect substitutability between bonds of different maturities observed in the data in a simple way.

Considering the utility function assumed in the SW model, after some algebra, the (linearized) consumption-based asset pricing equations can be written as

$$x_t = E_t x_{t+j} - \left(\frac{1-x_1}{\sigma_c}\right)\left[jr_t^{\{j\}} - \sum_{k=1}^{j} E_t\pi_{t+k} + j\varepsilon_t^{\{j\}}\right] + x_2\left(l_t - E_t l_{t+j}\right),\ for\, j=\ 1,\ 2,\ ...,,\ n, \tag{1}$$

where lower case variables denote the log-deviations of consumption (hours worked) from its

[10] As argued in Ireland (2004), there is a long-standing tradition of introducing additional disturbances into DSGE models in order to match the number of shocks with the number of time series used in estimation. This is because DSGE models typically introduce fewer shocks than observable variables, which means that certain combinations of endogenous variables are deterministic. If these combinations do not hold in the data, any approach that attempts to estimate the complete model will fail.

balanced-growth (steady-state) value or, alternatively, the deviations of the nominal interest rate, nominal yields, and the rate of inflation from their respective steady-state values. In particular, $r_t^{\{j\}}$ denotes the yield (written in deviations from its steady-state value) of a bond with a j-period maturity. The following notation is also used: $x_t = c_t - x_1 c_{t-1}$, $x_1 = \frac{h}{\gamma}$, $x_2 = \frac{(\sigma_c - 1)}{\sigma_c}\frac{WL}{C}$, where h and σ_c denote the habit formation and risk aversion parameters, and γ, W, L and C denote the balanced-growth rate and the steady-state values of the real wage, hours worked, and consumption, respectively. In particular, for $j = 1$ equation (1) can be written as equation (2) in Smets and Wouters (2007).

As in Eusepi and Preston (2011), the forward-looking behavior of both RE and AL agents can be captured by iterating the optimality condition (1), with j set to $j = 1$, j periods forward, and using the law of iterating projections, to obtain:

$$x_t = E_t x_{t+j} - \left(\frac{1 - x_1}{\sigma_c}\right) \sum_{k=1}^{j} E_t \left[r_{t+k-1}^{\{1\}} - E_t \pi_{t+k} + \varepsilon_{t+k-1}^{\{1\}} \right] + x_2 \left(l_t - E_t l_{t+j} \right). \tag{2}$$

Since equations (1) and (2) must hold in equilibrium, they imply the term structure of interest rates hypothesis:

$$r_t^{\{j\}} = \frac{1}{j} \sum_{k=0}^{j-1} E_t r_{t+k} + \xi_t^{\{j\}}, \tag{3}$$

where the supraindex $\{1\}$ on r_{t+k} has been removed for the sake of simplicity. Equation (3) states that the nominal yield of the j-period maturity bond, $r_t^{\{j\}}$, is equal to the average of the expectations for the short-term (1-period) nominal interest rate between periods t and $t + j - 1$, plus a term premium shock, $\xi_t^{\{j\}} = \left(\frac{1}{j}\sum_{k=1}^{j} E_t \varepsilon_{t+k-1}^{\{1\}}\right) - \varepsilon_t^{\{j\}}$, defined as the difference between the average of the expectations for the 1-period bond risk premium shocks between periods t and $t + j - 1$ and the risk premium shock associated with the j-period maturity bond. For the sake of simplicity, we assume that each term premium shock $\xi_t^{\{j\}}$

follows an AR(1) process:[11]

$$\xi_t^{\{j\}} = \rho^{\{j\}} \xi_{t-1}^{\{j\}} + \eta_t^{\{j\}}.$$

Furthermore, our empirical formulation below includes a constant term to capture the mean (steady-state value) of a yield.

The approach followed in our empirical analysis relies on the term structure hypothesis, equation (3) for $j = 4$, together with the set of log-linearized dynamic equations used in Slobodyan and Wouters (2012a) described in the Appendix, which among others includes the optimality condition (1) for $j = 1$.

According to equation (1), the term structure of interest rates can be understood as summarizing information on agents' beliefs about the future paths of inflation, consumption, and hours worked. More precisely, the nominal return associated with a j-period maturity bond must be consistent with the expected paths of inflation, consumption, and hours worked from period t until the maturity period $t + j$. Similarly, the term spread, $sp_t^{\{j\}} = r_t^{\{j\}} - r_t$ for $2 \leq j \leq n$, is clearly a forward-looking variable under both RE and AL since a (longer-term) interest rate, $r_t^{\{j\}}$, involves (according to the optimality equation (3)) the expectations of future realizations of the short-term nominal interest rate.Therefore, the term spread is also capturing the market's expectations about monetary policy. Moreover, the term spread is also driven by term premium innovations, $\eta_t^{\{j\}}$. As discussed below, the model estimated shows that term premium shocks become an important source of inflation fluctuations under AL. This finding stands in striking contrast to the absence of transmission of term premium shocks to the macroeconomy under RE. Furthermore, the consideration of term structure information in a DSGE model under AL contributes to the goal of disciplining expectations by

[11]This structure differs from the one considered by De Graeve, Emiris and Wouters (2009) in two aspects. First, they consider a measurement error in the term spread instead of a term premium shock. Second, they consider a time-varying inflation target in the monetary policy rule. The first difference is mainly a matter of semantics but the second may introduce an additional source of exogenous persistence. We choose to ignore this potential source of exogenous persistence for two main reasons. First, our empirical analysis shows that a time-varying inflation target is no longer needed to reproduce the actual aggregate persistence under AL. Second, this allows a more straightforward comparison with the Slobodyan and Wouters (2012a) model that helps to assess the importance of the AL expectations formation and the role of the term spread.

(i) characterizing agents' expectations beyond the one-period-ahead expectations considered in most DSGE models under AL; and (ii) using term structure information as observable in the estimation procedure.

3.3 The adaptive learning extension

How agents' beliefs are characterized becomes a crucial issue when one deviates from the RE hypothesis. This paper assumes small linear forecasting models that agents follow to form their expectations, the so-called "perceived law of motion" (PLM). Following Slobodyan and Wouters (2012a), the coefficients of the PLM are updated through a Kalman filter with the arrival of new information. Next, the small forecasting models are combined to form the expectation functions of the forward-looking variables of the model. The small forecasting models assumed are described below. A section in the Appendix briefly describes how learning coefficients are updated through a Kalman filter.

A PLM with term structure information

We adapt our extended SW model with term structure to the AL version of this model. One of the key ingredients of a model with AL is the way in which agents' expectations formation is characterized (i.e. the PLM of agents), so it is important to motivate the choice of the PLM. In an AL model with only 1-period-ahead expectations, Slobodyan and Wouters (2012a) suggest an AR(2) for each expectation formed at time t. They also include a time-varying intercept coefficient, which enables expectations to track down trend shifts in the data and changes in the inflation target.

In our DSGE model with term structure, we alternatively suggest a PLM motivated by (i) the interaction between term spreads and the expectations of consumption, inflation, and hours worked implicitly determined in the equilibrium condition (1); and (ii) the empirical evidence on the ability of term spreads to predict real economic activity and inflation (Estrella and Mishkin, 1997). At first sight, considering the whole term structure of interest rates to characterize AL might be seen as useful. However, considering term spreads associated with long-horizon bonds is rather challenging because it means having to define the whole set of expectations of the short-term nominal interest rate from the 1 -period horizon up to a long horizon. This cannot be done without imposing further restrictions on learning dynamics because, according to term structure equation (3), the number of parameters that define the PLM associated with these expectations dramatically increases with the number of expectations of the nominal short-term interest rate defined for alternative forecasting horizons, which in practice results in a curse of dimensionality problem. Furthermore, there is evidence (Mishkin, 1991) that at maturities longer than two quarters the term structure of interest rates does not help to anticipate future inflationary pressures.

For all these reasons, our empirical analysis focuses on the 1-year term spread (i.e. $sp_t^{\{4\}}$) because it implies a rather parsimonious AL model (thus, $j = 4$ in equation (3)). Moreover, following Slobodyan and Wouters (2008, 2012a), we allow agents to combine two forecasting models at the same time, track their forecasting performance, and use a variant of the Bayesian model averaging method to generate an aggregate forecast from the two forecasting models, which is used to characterize their decisions—see Slobodyan and Wouters (2008) for further details. The two forecasting models are

$$m_1 : \quad E_t y_{t+j} = \theta^{\{j\}}_{1,y,t-1} + \beta^{\{j\}}_{1,y,1,t-1} sp_t + \beta^{\{j\}}_{1,y,2,t-1} sp_{t-1},$$

$$m_2 : \quad E_t y_{t+j} = \theta^{\{j\}}_{2,y,t-1} + \beta^{\{j\}}_{2,y,1,t-1} sp_t,$$

where y_{t+j} stands for any variable that is being forecasted. Notice that these forecasting models are based only on current and lagged term spreads. In contrast to the forecasting models used in the related literature (e.g. Slobodyan and Wouters, 2012a,b), these two forecasting models depend only on real-time data.[12][13] Although these PLM are determined by actual information available to agents at the time when they formed their expectations, it is important to remark that the process of updating AL coefficients through the Kalman filter still depends on a few actual (revised) variables, such as aggregate consumption and inflation, since the AL coefficients are updated based on the forecast errors of these observables.[14] Indeed, this feature may explain the good model fit even though we consider small forecasting models based only on term structure information.

We also analyze two additional forecasting models.[15] The first model considers a third forecasting model given by an AR(2) as in Slobodyan and Wouters (2012a) in addition to models m_1 and m_2:

$$m_3 : \quad E_t y_{t+j} = \theta^{\{j\}}_{3,y,t-1} + \beta^{\{j\}}_{3,y,1,t-1} y_t + \beta^{\{j\}}_{3,y,2,t-1} y_{t-1},$$

whereas the second model includes the level of the 1-year rate, $r_t^{\{4\}}$, in m_3 instead of the AR(2) process, since its level is associated with inflation. That is,

$$m_3' : \quad E_t y_{t+j} = \theta^{\{j\}}_{3,y,t-1} + \beta^{\{j\}}_{3,y,1,t-1} r_t^{\{4\}}.$$

[12]

1. Readers might think that the second forecasting model adds nothing new to the first one. However, the simpler model m_2 may be useful in practice during episodes when the term spread is highly persistent. This is because the two regressors in model m_1 would be highly correlated, inducing a multicollinearity problem in those episodes which could be overcome by activating model m_2 instead.

[13]This timing assumption is used in Canova and Gambetti (2010), Slobodyan and Wouters (2012a,b), and Ormeño and Molnár (2015), among others. It is shared by RE models, but it may assume a richer information set than the set observed by agents when forming their forecasts in reality. Alternatively, one can assume that at time t agents only observe lagged values,sp_{t-1} . We investigate this alternative hypothesis below, and more deeply in Aguilar and Vázquez (2018).

[14]See equation (4a) in Slobodyan and Wouters (2012a). One possible way of overcoming this issue is to use additional vintage data, but this would inevitably increase the set of observables. Moreover, if the bulk of aggregate data revisions take place in the first revision, final revised data would be a good proxy of vintage revised data used in the updating procedure of AL coefficients.

[15]Moreover, a preliminary version of this paper, Aguilar and Vázquez (2015), investigated a wide range of PLM combining both revised data and term structure information in a single forecasting model. The overall fit of the DSGE model deteriorates when these complicated PLM are used.

As shown below, the inclusion of lagged variables or the level of the 1-year yield in the agents' forecasting model does not improve model fit once the term spreads are included.

4 Estimation results

In this section we first describe the data and the estimation approach. We then show the results of a few regression models for the observable forward-looking variables of the model to illustrate the relative predictive power of the 1-year term spread. Subsequently, the estimation results for the following four alternative DSGE models are discussed: (i) The SW model; (ii) the SW model with term spread; (iii) the SW model with AL suggested by Slobodyan and Wouters (2012a) (referred to hereafter as SlW); and (iv) the SW model with AL and term structure (hereafter SlWTS). Models (i) and (iii) are discussed and compared at length in Slobodyan and Wouters (2012a), whereas De Graeve, Emiris and Wouters (2009) discuss model variants of (i) and (ii). Therefore, our discussion focuses mainly on the interaction of AL expectations formation and the term structure of interest rates. This section also discusses the evolution of learning coefficients over time.

4.1 Data and the estimation approach

To facilitate comparison with the estimation results of Slobodyan and Wouters (2012a), the alternative models are estimated using U.S. data for a sample period running from 1966:1 until 2007:4. The set of observable variables is identical to theirs (i.e. the quarterly series of the inflation rate, the federal funds rate, the log of hours worked, and the quarterly log differences of real consumption, real investment, real wages, and real GDP) with the addition of the 1-year Treasury constant maturity yield. GDP, consumption, investment, and hours worked are measured in per-working age population terms. The measurement equation is

$$X_t = \begin{bmatrix} dlGDP_t \\ dlCONS_t \\ dlINV_t \\ dlWAG_t \\ lHours_t \\ dlP_t \\ FEDFUNDS_t \\ One-year\,TB\,yield_t \end{bmatrix} = \begin{bmatrix} \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \bar{l} \\ \overline{\pi} \\ \overline{r} \\ \overline{r}^{\{4\}} \end{bmatrix} + \begin{bmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - i_{t-1} \\ w_t - w_{t-1} \\ l_t \\ \pi_t \\ r_t \\ r_t^{\{4\}} \end{bmatrix},$$

where l and dl denote the log and the log difference, respectively. $\overline{\gamma} = 100(\gamma - 1)$ is the common quarterly trend growth rate for real GDP, real consumption, real investment and real wages, which are the variables that feature a long-run trend. $\bar{l}$, $\overline{\pi}$, $\overline{r}$ and $\overline{r}^{\{4\}}$ are the steady-state levels of hours worked, inflation, the federal funds rate and the 1-year (four-quarter) constant maturity Treasury yield, respectively.

The procedure for estimation under AL is essentially the same as in Slobodyan and Wouters (2012a), except that it is modified to take into account of the fact that agents are using both multiperiod forecasting and a set of alternative forecasting models based only on term spread information. The estimation approach follows a two-step Bayesian estimation procedure. First, the log posterior function is maximized by combining prior information on the parameters with the likelihood of the data. The prior assumptions are exactly the same as in Slobodyan and Wouters (2012a). Moreover, we consider rather loose priors for the parameters that characterize the 1-year yield dynamics. The second step implements the Metropolis-Hastings algorithm, which runs a massive sequence of draws of all the possible realizations for each parameter in order to draw a picture of the posterior distribution.[16]

[16]The RE versions of the DSGE models are estimated using standard Dynare routines whereas the AL versions of the models use the codes kindly provided by Sergey Slobodyan and Raf Wouters with a few modifications to accommodate the presence of the term spread in the PLM.

4.2 The predictive power of the term spread

Before discussing the estimates found in the Bayesian estimation of the DSGE model with AL, we illustrate the relative predictive power of the 1-year term spread by running five regression models for seven observable forward-looking variables of the model.

The first regression model includes the first lag of the dependent variable. The second model includes the first lags of the growth rates of consumption, investment, and the real wage together with the first lags of hours worked, inflation, and the short-term rate. The third model includes the first lag of the term spread. The fourth model includes the first lag of the short-term interest rate. Finally, the fifth model includes the first lag of the one-year yield. In addition, all regression models include a constant along with linear and quadratic trends to capture fluctuations in low-frequency components. The first two models include revised data in general, while the other three use either the term spread or interest rates, which are observed in real time.[17]

[17] This regression analysis should be viewed as a simple illustration of the predictive power of the term spread to predict the observable forward-looking variables of the model under AL because these variables enter the model measured either in deviations from the balance growth path or in deviations from their respective steady-state values (neither is actually observed by the econometrician) instead of the growth rates or the levels used as dependent variables in the regression analysis. However, we consider the growth rates of consumption, investment, and the real wage instead of any alternative definition of their cyclical components, such as their linear detrended components, for two main reasons: First, the growth rates of these variables are the observables that enter into the measurement equation. Second, as argued below, the time-varying intercept of the PLMs captures the low-frequency (i.e. long-run) components of the data, whereas the remaining variables in the PLM explain mainly the high frequency components described by the growth rates.

Table 1. Adjusted R^2 of alternative regression models

	Regression models				
Dependent variables	(1)	(2)	(3)	(4)	(5)
consumption growth	0.03	0.18	0.21	0.14	0.08
investment growth	0.27	0.39	0.23	0.22	0.15
real wage growth	0.02	0.03	0.01	0.01	0.01
hours worked	0.95	0.97	0.11	0.13	0.14
inflation	0.75	0.76	0.47	0.48	0.44
short-term interest rate	0.90	0.92	0.55	0.90	0.88
short rate 4-quarters-ahead	0.51	0.58	0.38	0.51	0.53

Notes: All regression models include a constant, and linear and quadratic trends to capture fluctuations in low-frequency components. In addition, model (1) includes the first lag of the dependent variable; model (2) includes the first lags of the growth rates of consumption, investment and the real wage together with the first lags of hours worked, inflation and the short-term rate; model (3) includes the first lag of the term spread; model (4) includes the first lag of the short-term interest rate; and model (5) includes the first lag of the one-year yield.

Table 1 shows the (adjusted) R^2 for the alternative regression models. The term spread regression model does at least as well as the other models in predicting the growth rates of consumption, investment and the real wage one quarter ahead. However, it does much worse than the first two (and similarly to the last two) in predicting one-quarter ahead inflation and hours worked, and the fourth-quarter-ahead short term interest rate. Finally, the term spread does a reasonable job in predicting short-term interest rates. Overall this regression analysis shows that the 1-year term spread can be helpful in forecasting many of the forward-looking variables of the model in real time.

4.3 Main empirical findings

Table 2 shows the estimation results associated with the four alternative DSGE models estimated divided into two panels. Panel A shows the marginal likelihood and the structural parameter estimates, while Panel B shows the estimates of the parameters that describe shock processes. In line with the findings of Slobodyan and Wouters (2012a), the marginal likelihood of the AL model without term structure using their PLM featured by an AR(2), SlW, is larger (-960.22) than that associated with the RE model, SW model (-973.76). However, when term structure information is included in the two models, the RE version performs similarly (-864.89) to our AL version (-867.97), which features small forecasting models based only on term structure information.[18] This finding might come as a surprise because the forecasting models assumed can be viewed on the one hand as rather realistic given that they ignore information on revised macroeconomic data which is not available to economic agents on forming their expectations; but on the other hand they look rather restrictive in that they take into account only term structure information.

[18]As pointed out by Del Negro and Eusepi (2011, p.2116), a small difference of 5 points or less can be overturned by choosing a slightly different set of priors.

Table 2. Panel A: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posteriors											
				SW			SW spread			SlW			SlWTS		
Log-lik.				-973.76			-864.89			-960.22			-867.97		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
φ	N	4.00	1.50	5.96	4.21	7.63	6.06	4.33	7.89	3.34	1.88	3.87	1.72	1.14	2.32
h	B	0.70	0.10	0.79	0.74	0.86	0.77	0.69	0.83	0.68	0.63	0.75	0.77	0.72	0.83
σ_c	N	1.50	0.37	1.22	1.04	1.34	1.22	1.02	1.45	1.53	1.19	1.63	0.56	0.46	0.63
σ_l	N	2.00	0.75	1.50	0.64	2.32	1.82	0.89	2.53	1.74	1.02	2.60	2.12	1.37	2.78
ξ_p	B	0.50	0.10	0.70	0.62	0.80	0.74	0.57	0.88	0.64	0.59	0.69	0.42	0.35	0.49
ξ_w	B	0.50	0.10	0.71	0.62	0.78	0.86	0.81	0.91	0.82	0.76	0.85	0.66	0.58	0.72
ι_w	B	0.50	0.15	0.51	0.30	0.72	0.24	0.10	0.42	0.18	0.07	0.26	0.45	0.25	0.64
ι_p	B	0.50	0.15	0.25	0.10	0.38	0.27	0.10	0.39	0.27	0.11	0.39	0.17	0.07	0.27
ψ	B	0.50	0.15	0.55	0.36	0.72	0.53	0.34	0.74	0.50	0.31	0.71	0.58	0.38	0.78
Φ	N	1.25	0.12	1.62	1.48	1.73	1.57	1.41	1.68	1.58	1.45	1.73	1.38	1.26	1.49
r_π	N	1.50	0.25	1.98	1.71	2.25	1.80	1.15	2.09	1.74	1.38	2.04	1.69	1.45	1.99
ρ_r	B	0.75	0.10	0.84	0.80	0.87	0.87	0.83	0.89	0.88	0.85	0.91	0.80	0.76	0.85
r_y	N	0.12	0.05	0.10	0.05	0.13	0.11	0.06	0.15	0.13	0.07	0.18	0.16	0.10	0.22
$r_{\Delta y}$	N	0.12	0.05	0.15	0.12	0.18	0.17	0.14	0.20	0.13	0.10	0.16	0.20	0.16	0.24
r_{sp}	N	0.12	0.05	-	-	-	-	-	-	-	-	-	0.14	0.06	0.21
π	G	0.62	0.10	0.67	0.51	0.82	0.71	0.54	0.88	0.63	0.53	0.74	0.69	0.53	0.86
β	G	0.25	0.10	0.20	0.08	0.28	0.20	0.10	0.30	0.18	0.09	0.27	0.24	0.10	0.37
l	N	0.00	2.00	1.05	-0.5	2.14	0.59	-0.72	2.03	1.10	-0.76	1.96	0.01	-1.33	1.21
γ	N	0.40	0.10	0.40	0.38	0.42	0.41	0.39	0.42	0.40	0.36	0.43	0.43	0.40	0.45
$\bar{r}^{\{4\}}$	N	1.00	0.50	-	-	-	1.38	1.19	1.58				1.23	0.77	1.70
α	N	0.30	0.05	0.19	0.16	0.22	0.18	0.15	0.20	0.17	0.13	0.19	0.17	0.13	0.19
ρ	B	0.50	0.28	-	-	-	-	-	-	0.97	0.96	0.98	0.84	0.79	0.98

Table 2. Panel B: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posterior											
				SW			SW spread			SlW			SlWTS		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
σ_a	IG	0.10	2.00	0.44	0.39	0.48	0.45	0.40	0.49	0.46	0.41	0.49	0.49	0.45	0.55
σ_b	IG	0.10	2.00	0.25	0.21	0.29	0.24	0.20	0.28	0.15	0.14	0.20	0.89	0.78	0.99
σ_g	IG	0.10	2.00	0.52	0.47	0.57	0.51	0.48	0.56	0.50	0.46	0.56	0.50	0.45	0.55
σ_i	IG	0.10	2.00	0.44	0.37	0.52	0.44	0.36	0.52	0.46	0.42	0.52	1.57	1.41	1.72
σ_R	IG	0.10	2.00	0.22	0.20	0.24	0.22	0.20	0.24	0.21	0.19	0.23	0.24	0.21	0.26
σ_p	IG	0.10	2.00	0.14	0.11	0.17	0.17	0.14	0.19	0.15	0.13	0.16	0.30	0.27	0.33
σ_w	IG	0.10	2.00	0.23	0.20	0.27	0.30	0.27	0.33	0.21	0.18	0.22	0.56	0.51	0.61
$\sigma^{\{4\}}$	IG	0.10	2.00	-	-	-	0.10	0.09	0.11				0.19	0.17	0.22
ρ_a	B	0.50	0.20	0.93	0.89	0.96	0.93	0.90	0.97	0.98	0.98	0.99	0.96	0.92	0.99
ρ_b	B	0.50	0.20	0.17	0.04	0.26	0.26	0.15	0.40	0.43	0.29	0.55	0.91	0.87	0.95
ρ_g	B	0.50	0.20	0.98	0.97	0.99	0.98	0.97	0.99	0.96	0.95	0.98	0.98	0.96	0.99
ρ_i	B	0.50	0.20	0.70	0.59	0.79	0.69	0.59	0.79	0.50	0.37	0.57	0.96	0.93	0.98
ρ_R	B	0.50	0.20	0.08	0.01	0.14	0.06	0.01	0.11	0.09	0.01	0.16	0.15	0.05	0.27
ρ_p	B	0.50	0.20	0.84	0.74	0.95	0.85	0.75	0.99	0.32	0.06	0.57	0.87	0.79	0.94
ρ_w	B	0.50	0.20	0.97	0.96	0.99	0.95	0.93	0.98	0.54	0.32	0.80	0.97	0.96	0.98
$\rho^{\{4\}}$	B	0.50	0.20	-	-	-	0.84	0.77	0.91	-	-	-	0.94	0.90	0.99
μ_p	B	0.50	0.20	0.67	0.49	0.85	0.75	0.61	0.90	0.47	0.29	0.67	0.41	0.26	0.53
μ_w	B	0.50	0.20	0.87	0.82	0.94	0.98	0.86	0.99	0.43	0.11	0.70	0.26	0.15	0.36
ρ_{ga}	B	0.50	0.20	0.51	0.37	0.65	0.52	0.39	0.67	0.17	0.13	0.19	0.54	0.41	0.66

Notes: The labels N, B G and IG denote normal, beta, gamma and inverse gamma prior distribution, respectively.

In order to motivate the use of PLMs based only on term spreads, we estimated two additional models as discussed above. Thus, we firstly analyzed a forecasting model in which lagged (revised) endogenous variables are included in the PLMs (model m_3) in addition to the forecasting models based on term spreads considered in the baseline case. Second, we considered a forecasting model in which the level of the 1-year yield is included in the PLMs (model m_3') in addition to the forecasting models based on term spreads. Table 3 shows that including either lagged endogenous variables or the level of the 1-year rate does not improve model fit once the term spread is included in the PLMs (models m_1 and m_2). As argued above, the good performance of the PLM including term spreads is due to the way in which AL coefficients are updated through the Kalman filter. Namely, the learning coefficients are updated based on the forecast errors and those forecast errors depend on actual values of forward-looking variables, such as aggregate consumption and inflation. In this way, a few learning coefficients are updated using actual data. In particular, the time-varying intercept is able to capture the low frequency movement of inflation as shown in Figure 1 below. Moreover, in reality term spreads might be capturing other information that agents have about the course of the economy not otherwise present in the model. As a consequence, the reasonable performance of the AL model with term structure may be due to the information conveyed by term spreads over and above that captured in the model.

As also shown in Table 3, a simple PLM which includes only a time-varying intercept fits the data much worse than the PLMs based on term structure information. Nonetheless, such a simple PLM based only on a time-varying intercept fits the data better than PLMs that include many more variables analyzed in Section 6 below. In particular, we show that the marginal likelihood decreases rather dramatically if the small forecasting models follow the AR(2) structure assumed in Slobodyan and Wouters (2012a). The reason is rather simple: Since many variables in the model largely comove, including a few of those variables in the same forecasting model may induce (multi-) collinearity problems. Why might multicollinearity be a problem? The chances of converging to a local maximum in this scenario increase because the likelihood function becomes flat in some regions of the parameter space under multicollinearity. These considerations certainly make a compelling case for the use of a combination---i.e. model averaging--- of small forecasting models instead of a single forecasting model with many explanatory variables. The PLMs assumed in this paper are very simplistic, but the bottom line is that the time-varying intercepts of the PLMs capture the long-run features of data (e.g. inflation) well whereas the time-varying coefficients associated with term spreads seem to capture short-run movements (growth rates of consumption and investment) relatively well as shown in Table 1.

Table 3. Model's fit for alternative PLMs

	Baseline	Adding AR(2)	Adding 1-year yield	Only intercept
log-likelihood	-865.20	-865.41	-866.42	-910.83

Our estimation results also suggest that the fitting of the short-term yield curve summarized by the 1-year yield becomes somewhat more challenging for the AL model than for the RE model. That is, the increase in the log marginal likelihood when the 1-year yield is considered as observable in the estimation procedure is much larger under RE, at 108.87 ($= -864.89 - (-973.76)$), than under AL, at 92.25 ($= -867.97 - (-960.22)$).[19]

Table 2 shows that in general many parameter estimates are rather robust across models, with a few important differences. These differences are discussed in two parts: First, the differences between the SW model under RE and the SW model under AL (i.e. SW model versus SlW model) are discussed to assess the contribution of AL. Second, the differences between the SW model and the SW model with term spread, both under AL (i.e. SlW model versus SlWTS), are studied to determine the contribution of the term spread when interacting with AL.[20]

[19]See Del Negro and Eusepi (2011) for a discussion of the econometric framework, based on log marginal likelihood differences, for assessing how a model estimated to fit a benchmark set of time series (in our case, the set of observables in Slobodyan and Wouters, 2012a) performs on fitting an additional time series (the 1-year yield).

[20]We also consider the 1-year Treasury constant maturity yield to estimate the SW model under RE. Introducing the 1-year yield in the SW model barely changes parameter estimates with a few exceptions.

SW model versus SlW

As found by Slobodyan and Wouters (2012a), considering AL instead of the RE assumption in the SW model estimated greatly reduces the importance of both exogenous persistence due to price markup and wage markup shocks and endogenous persistence induced by habit formation, the elasticity of the cost of adjusting capital, and wage indexation. The intuition of these findings is rather simple: AL dynamics introduce a major channel of endogenous persistence that is ignored when the RE assumption is considered. As a consequence, a few sources of persistence under the RE assumption are less significant under AL in reproducing the observed persistence in most macroeconomic variables.

SlW model versus SlWTS

The introduction of the term spread into the PLM results in many more changes than those introduced by the single-step extension of the SW model with AL analyzed above. On the one hand, a few sources of endogenous persistence are less important. Thus, the estimates of the inverse of consumption intertemporal elasticity, σ_c, and the elasticity of the cost of adjusting capital, φ, are much smaller under AL with term structure (0.56 and 1.72, respectively) than without (1.53 and 3.34, respectively).[21] Similarly, the price and wage probabilities (ξ_p and ξ_w) and the price indexation parameter, ι_p, are much smaller under AL with term structure than without term structure, whereas the opposite occurs for the wage indexation parameter, ι_w. We find contrasting results as regards exogenous sources of price and wage markup

---

The Calvo wage probability estimate, ξ_w, increases from 0.71 to 0.86 when the term spread is considered. The opposite occurs for the wage indexation parameter, ι_w, that goes from 0.51 to 0.24. These results suggest that the relative importance of endogenous sources (compared to exogenous sources) in explaining price and wage persistence increases when the RE model is extended with the short-term yield dynamics. Moreover, the inverse of the Frisch elasticity of labor supply, σ_l, the volatility of the innovation, and the moving average coefficient associated with the wage shock, σ_w and μ_w, and the persistence of the risk premium shock, ρ_b, increase slightly.

21

1. As emphasized in Smets and Wouters (2007), a lower elasticity of the cost of adjusting capital increases the sensitivity of investment to the real value of the existing capital stock, Tobin's q (see equation (5) in the Appendix).

persistence. Thus, the autoregressive parameters of the ARMA processes that characterize price and wage markup shocks (ρ_p and ρ_w) are higher in the AL model with term structure (and similar to those estimated in the RE models) than in the model without term structure. However, the moving average parameters of the ARMA processes that characterize price and wage markup shocks (μ_p and μ_w) are lower in the AL model with term structure than in the AL model without term structure. In short, we observe that AL under the multi-period forecasting hypothesis associated with the inclusion of term structure information results in lower estimates of the parameters that characterize important sources of endogenous persistence than those implied by either the AL model without term structure information or the RE model. These results are in line with those found in Eusepi and Preston (2011) in the sense that AL with multi-period forecasting takes over in explaining aggregate persistence while other sources become less important.

On the other hand, it is important to recall that the term spread is a forward-looking variable. This means that learning dynamics endowed with term spread information are less sluggish. Thus, the estimated persistence of belief coefficients, ρ, is lower when the term spread in the SlWTS model (0.84) is considered than in the SlW model (0.97). As a consequence of the much faster adjustment of belief coefficients, the estimates of most parameters that capture exogenous persistent dynamics in the SlWTS (i.e. ρ_b, ρ_i, ρ_p and ρ_w) are much higher, so they mimic actual data persistence. Moreover, a lower value of ρ means not only less sluggish learning dynamics but also that the belief coefficients themselves are much less volatile. When the belief coefficients are less volatile the forecasting models have less chance of hitting the projection facility boundary (i.e. the beliefs react less to the same forecasting errors). Hitting the projection facility boundary is costly in terms of likelihood, so this is a valuable feature which enables the initial values of the beliefs to be initialized closer to the unit root boundary. Since we follow Slobodyan and Wouters (2012a,b) and initialize the beliefs by calculating them at the RE equilibrium implied by the estimated parameters, larger values of the estimated parameters that feature exogenous persistence are also consistent with the smaller estimated value of ρ.[22]

Our results also show that the risk premium and investment disturbances—and to a lesser extent markup disturbances—are more volatile under AL based on term structure information. The sensitivity of the estimated shock process parameters to the specific perceived law of motion considered may suggest that these disturbances can be viewed as simply wedges (i.e. reduced-form shocks) rather than truly structural shocks as argued in Chari, Kehoe and McGrattan (2009).[23]

4.4 Analysis of the PLM

Figure 1 shows the trend in the PLM coefficients for inflation and consumption. We focus on these two forward-looking variables because they have observable counterparts. The time-varying intercept of the PLM of inflation (consumption) shows how agents' perceptions of steady-state inflation (balanced-growth consumption) vary over time. Thus, the intercept of inflation expectations captures the rise of inflation expectations during the 60's and 70's and the rapid fall of inflation expectations from the early 80's onward. The intercept of consumption captures the fall in long-term consumption expectations from the early 70's to the mid-80's, capturing the economic recessions around the oil price shocks, followed by a steady rise in long-term consumption expectations up to the early 2000's.

[22] We thank Sergey Slobodyan for pointing this out.

[23] This sensitivity of risk premium shock parameters is in line with the different estimates reported for these parameters in Smets and Wouters (2007) and Galí, Smets and Wouters (2012), which were obtained using a similar sample period.

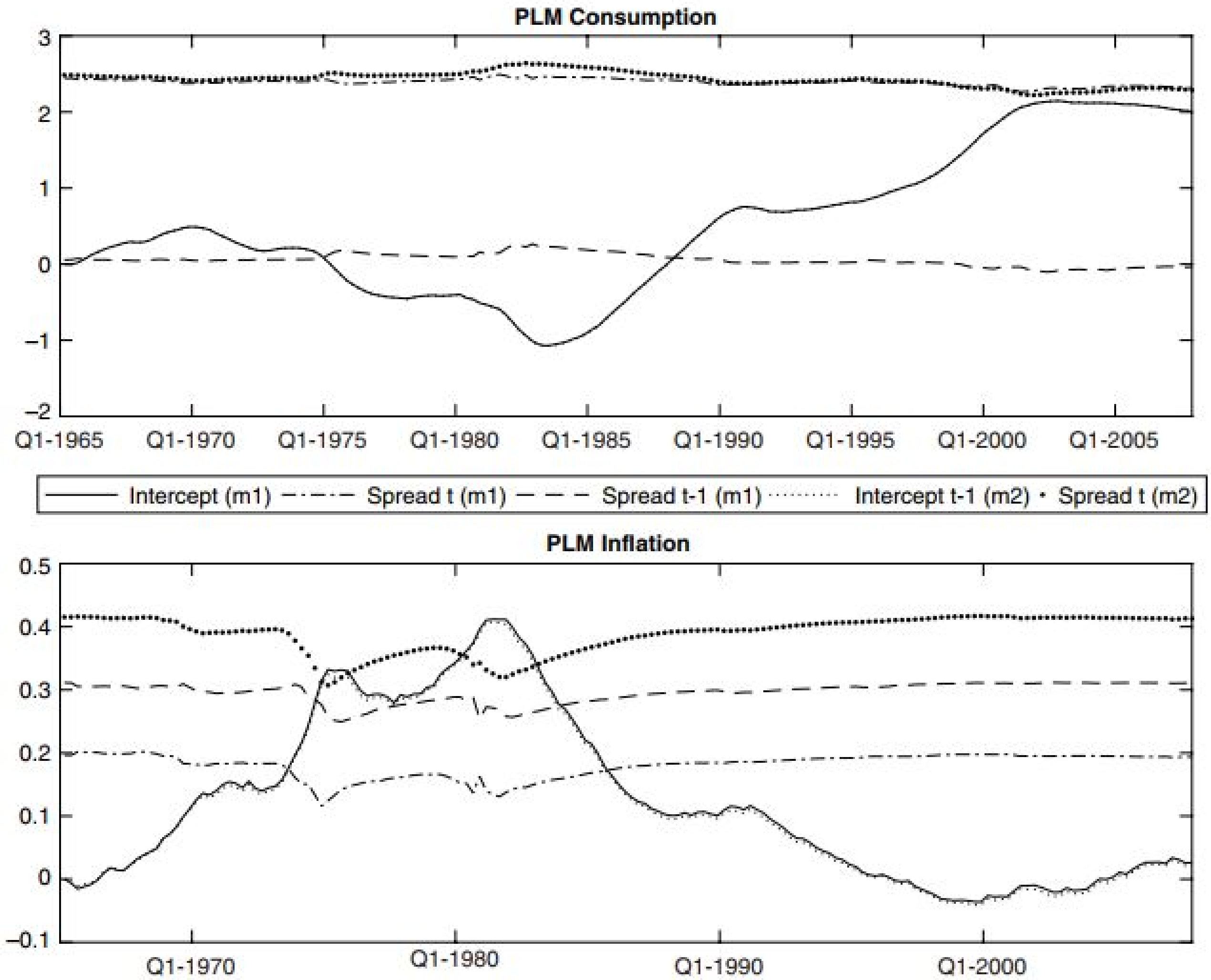


Figure 1. PLM coefficients of inflation and consumption

In contrast to the time-varying intercept coefficients, the term spread coefficients in the two PLM are fairly stable. Moreover, the term-spread coefficients associated with these two PLM are positive, indicating that a higher 1-year bond yield today anticipates tighter financial conditions in the future (say 4-quarters ahead), which results in higher one-quarter ahead expectations of inflation and consumption because agents substitute current consumption for future consumption. Notice that the sum of the term-spread coefficients associated with the PLM of consumption is roughly five times larger than that associated with the PLM of inflation, which results in much larger swings in consumption expectations than in inflation expectations due to term spread changes. This greater sensitivity of consumption expectations (relative to inflation expectations) to term spread variations is consistent with the relative size of the impulse responses of consumption and inflation to a term spread innovation shown below. Also, notice that the small forecasting model m_1 characterizes inflation expectations differently because both contemporaneous and lagged spreads have non-trivial coefficients in the agents' forecasting function, while consumption expectations are described by the simpler m_2 forecasting model. This empirical finding further motivates the combination of two alternative small forecasting models to describe all forward-looking variables of the model.

5 The transmission of shocks

As shown by Slobodyan and Wouters (2012a,b), the transmission of structural shocks is crucially determined by the way in which agents form their expectations. Therefore, it is important to analyze how impulse response functions (IRF) shift over time driven by changes in the updating belief coefficients. Our IRF analysis focuses on the estimated time-varying responses of a selected group of real (i.e. output and consumption) and nominal (inflation and the short-term interest rate) variables to a term premium innovation. As emphasized above, the term spread is a forward-looking variable. Thus, the responses of output and inflation to term premium innovations illustrate how term structure innovations anticipate movements in these variables.[24]

[24] Following Slobodyan and Wouters (2012a), the IRF are computed using the fixed belief coefficients obtained using the information available at each point in time, but then ignoring the updating of those beliefs driven by the shock. Therefore, these IRF might underestimate the size and persistence of actual responses.

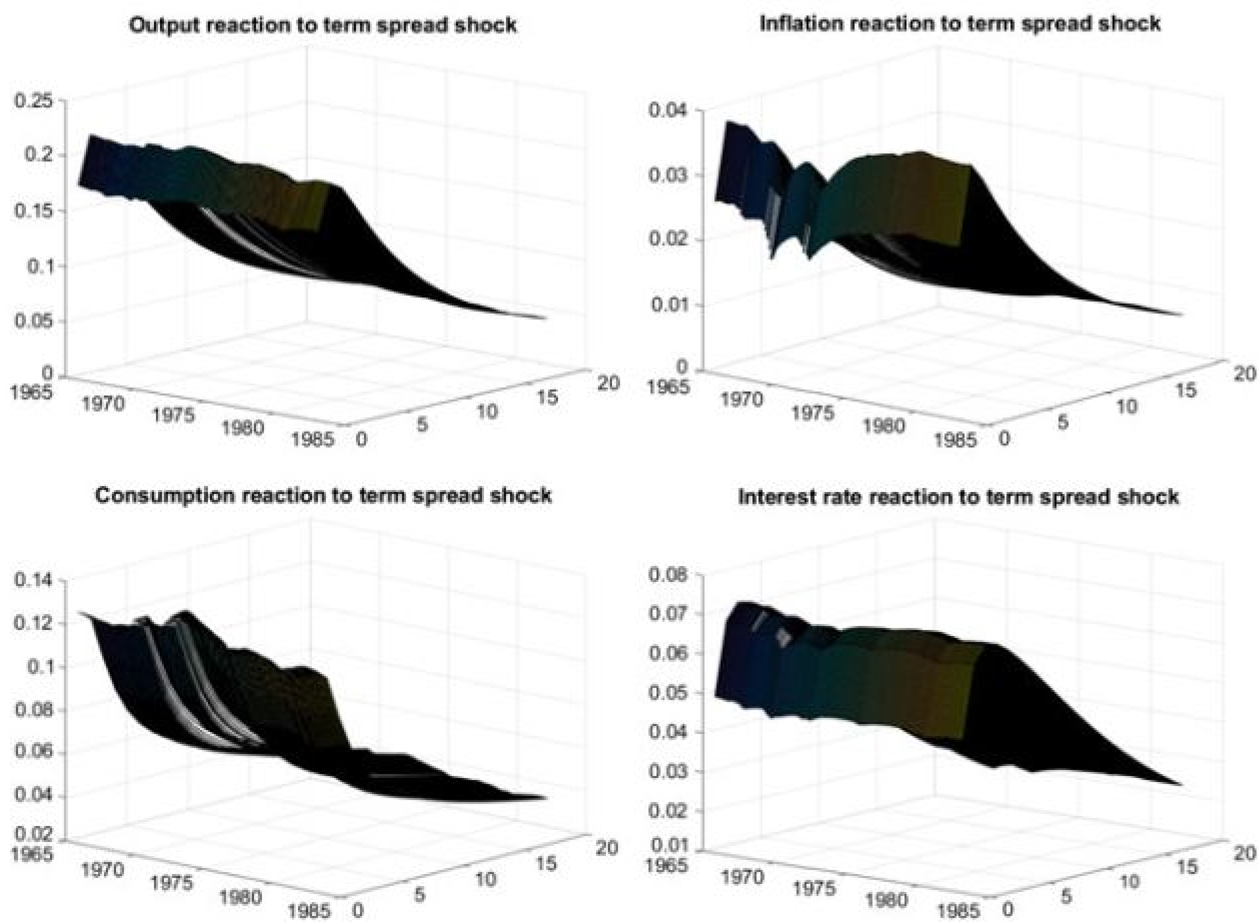


Figure 2. IRF to a term premium shock

5.1 Impulse responses to a term premium shock

The introduction of AL extended with term structure information permits a feedback from the term structure to the macroeconomy through the learning dynamics that is missing under

RE. Figure 2 shows the time-varying IRF of output, consumption, inflation, and the nominal interest rate to a term premium innovation under AL. The stability of term-spread learning coefficients associated with the PLM featured only by the term spread shown above means that the IRF change modestly over time, with more prominent changes in the responses during the 70's and early 80's. A positive innovation in the term premium shock increases the 1-year yield relative to the (short-term) federal funds rate, which brings forward consumption and investment decisions, resulting in higher current economic activity (output and consumption), inflation, and (short-term) nominal interest rate. The impulse responses of all variables are hump-shaped, capturing a different transmission mechanism (relative to RE). This hump-shaped feature is more pronounced in the nominal variables (inflation and nominal interest rate) than in the real variables (output and consumption).

5.2 Variance-decomposition analysis

Table 4 shows the variance decomposition of output, consumption, inflation, and the short term nominal interest rate for two alternative forecasting horizons. Risk premium shocks explain a large proportion of output and consumption variability at short-term forecasting horizons, whereas both productivity and exogenous spending significantly increase their proportions as the forecasting horizon increases.

Monetary policy shocks explain a large proportion of inflation variability at short-term forecasting horizons, but term premium shocks account for an increasing proportion at long-term forecasting horizons (32.84%). This last feature also shows up when the variability of the short-term nominal interest rate is analyzed. Thus, 28.85% of the short rate long-run variability is explained by the term premium shock. Interestingly, price and wage markup shocks play a small role in explaining inflation variability. This is in contrast to the variance decompositions reported in Smets and Wouters (2007) and Slobodyan and Wouters (2012a), where these two shocks explain a sizable share of inflation variability.

Table 4. Variance decomposition

	1-year horizon				10-year horizon			
	y	c	π	r	y	c	π	r
Productivity	22.82	7.54	6.85	7.74	31.61	7.87	10.33	11.48
Risk premium	51.40	74.70	0.11	18.61	35.60	59.74	0.09	9.37
Exogenous spending	13.25	10.26	0.08	4.42	23.07	23.26	0.06	3.47
Investment technology	3.34	2.02	1.77	44.93	1.42	0.99	1.45	21.85
Monetary policy	1.46	0.74	76.80	15.52	1.47	0.96	51.36	18.19
Price markup	5.06	3.56	0.06	1.61	5.09	5.21	0.90	1.64
Wage markup	2.62	1.08	2.45	3.55	1.45	0.73	2.96	5.15
Term premium	0.06	0.08	11.87	3.62	0.28	1.25	32.84	28.85

In short, our results suggest that term premium innovations together with monetary innovations take over markup shocks in explaining nominal variability. However, their role in real variables is actually small, with the risk premium shock playing a prominent role, especially in explaining consumption variability.

6 Robustness analysis

This section studies the robustness of the estimation results of the SlWTS model using four alternative specifications for the PLM across two dimensions: in-sample fit and parameter estimates. The first alternative specification (labeled "I") uses the PLM suggested by Slobodyan and Wouters (2012a) (i.e. an AR(2)). The second alternative formulation (labeled "II") considers the minimum set of state variables. The third specification (labeled "III") includes only the lagged term spread in all PLM. Finally, the fourth specification (labeled "IV") considers the baseline learning specification, but with the belief coefficient, ρ, restricted to zero. The baseline formulation and the last two use only term spread information observed in real time, while the other two use revised data. Table 5 shows the estimation results for the five alternative specifications of the PLM studied.[25] The marginal likelihood evaluated at the mode clearly favors the baseline specification. In particular, notice that the (posterior) marginal likelihood for the Slobodyan and Wouters estimation is -960.22 in Table 2, whereas the marginal likelihood computed at the mode is -958.62. When the 1-year yield is included in the set of observables, the marginal likelihood for the Slobodyan and Wouters estimation is -952.45 at the mode as reported in Table 5. This small improvement in model fit when adding the 1-year yield as observable in the Slobodyan and Wouters estimation is in contrast to the large fit improvement found in our model: 6.17 ($= -952.45 - (-958.62)$) versus 101.27 ($= -865.20 - (-966.47)$), where all these values are computed at the mode. This finding suggests that the fitting of the 1-year yield becomes much more challenging for the forecasting models of Slobodyan and Wouters than for the forecasting models based on term spreads. Moreover, the use of the lagged term spread instead of the current term spread results in a large deterioration in the fit of the model. The conclusions are similar for the other three formulations. The deterioration becomes much larger when the learning coefficient, ρ, is restricted to zero.

[25] All models in Table 5 are estimated using the same set of observables, which includes the seven observables in Slobodyan and Wouters (2012a) and the 1-year yield.

Table 5. Panel A: Robustness analysis of alternative PLM: structural parameters

	Baseline		I		II		III		IV	
Log. Lik.	-865.20		-952.45		-952.01		-933.63		-1155.90	
	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.
φ	1.60	0.38	4.46	0.18	4.74	0.24	1.34	0.38	1.02	0.02
h	0.78	0.04	0.65	0.11	0.81	0.07	0.73	0.07	0.62	0.02
σ_c	0.50	0.11	1.37	0.15	1.20	0.14	0.53	0.07	0.70	0.02
σ_l	2.16	0.48	1.55	0.10	2.05	0.19	2.00	0.51	1.41	0.01
ξ_p	0.44	0.09	0.48	0.13	0.70	0.14	0.44	0.10	0.43	0.02
ξ_w	0.63	0.04	0.58	0.12	0.57	0.14	0.62	0.16	0.91	0.02
ι_w	0.38	0.17	0.18	0.13	0.43	0.21	0.37	0.24	0.31	0.02
ι_p	0.16	0.26	0.39	0.14	0.55	0.17	0.32	0.39	0.23	0.02
ψ	0.49	0.24	0.04	0.10	0.11	0.17	0.34	0.49	0.87	0.03
Φ	1.36	0.08	1.51	0.13	1.57	0.12	1.44	0.11	1.35	0.03
r_π	1.73	0.17	1.84	0.11	1.86	0.19	1.72	0.35	2.05	0.02
ρ_r	0.81	0.04	0.86	0.08	0.89	0.05	0.63	0.04	0.88	0.02
r_y	0.13	0.20	0.03	0.14	0.14	0.20	-	-	0.11	0.02
$r_{\Delta y}$	0.20	0.08	0.15	0.11	0.15	0.15	-	-	0.35	0.02
r_{sp}	0.14	0.09	0.07	0.11	0.15	0.17	0.14	0.47	0.11	0.02
π	0.65	0.10	0.84	0.08	0.61	0.19	0.63	0.16	0.79	0.02
β	0.22	0.08	0.27	0.15	0.17	0.16	0.34	0.22	0.35	0.02
l	-0.32	0.34	-0.08	0.18	-1.02	0.21	-1.43	0.52	-1.02	0.02
γ	0.43	0.02	0.39	0.09	0.40	0.05	0.41	0.04	0.44	0.03
$\bar{r}^{\{4\}}$	1.30	0.16	1.99	0.18	2.44	0.24	1.49	0.35	1.43	0.02
α	0.15	0.05	0.18	0.11	0.19	0.16	0.15	0.06	0.23	0.03
ρ	0.95	0.03	0.99	0.03	0.99	0.03	0.97	0.01	0.0	-

Table 5. Panel B: Robustness analysis of alternative PLM: shock process parameters

	Baseline		I		II		III		IV	
	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.
σ_a	0.50	0.05	0.44	0.11	0.47	0.12	0.48	0.06	0.50	0.02
σ_b	0.90	0.08	0.22	0.13	0.31	0.12	0.79	0.08	4.69	0.02
σ_g	0.49	0.03	0.50	0.14	0.51	0.10	0.50	0.07	0.50	0.02
σ_i	1.52	0.10	0.46	0.13	0.94	0.14	1.54	0.09	1.75	0.02
σ_R	0.23	0.02	0.23	0.14	0.22	0.14	0.24	0.07	0.24	0.02
σ_p	0.29	0.05	0.16	0.10	0.16	0.16	0.27	0.12	0.32	0.02
σ_w	0.55	0.04	0.61	0.14	0.44	0.09	0.55	0.06	0.56	0.01
$\sigma^{\{4\}}$	0.19	0.02	0.15	0.10	0.19	0.09	0.15	0.02	0.16	0.02
ρ_a	0.95	0.02	0.88	0.12	0.97	0.03	0.94	0.04	0.98	0.01
ρ_b	0.90	0.02	0.56	0.14	0.23	0.17	0.96	0.02	0.76	0.02
ρ_g	0.98	0.01	0.98	0.15	0.99	0.01	0.98	0.01	0.97	0.03
ρ_i	0.96	0.02	0.58	0.14	0.59	0.21	0.95	0.03	0.93	0.02
ρ_R	0.12	0.23	0.23	0.16	0.09	0.18	0.59	0.18	0.12	0.02
ρ_p	0.85	0.06	0.84	0.14	0.59	0.18	0.87	0.10	0.76	0.03
ρ_w	0.97	0.10	0.88	0.09	0.91	0.09	0.98	0.03	0.98	0.02
$\rho^{\{4\}}$	0.92	0.04	0.95	0.07	0.96	0.03	0.93	0.06	0.97	0.02
μ_p	0.41	0.11	0.55	0.13	0.59	0.14	0.57	0.16	0.39	0.02
μ_w	0.29	0.16	0.24	0.15	0.44	0.16	0.28	0.26	0.07	0.01
ρ_{ga}	0.54	0.09	0.62	0.12	0.57	0.15	0.52	0.25	0.53	0.02

Notes: Log-likelihood values evaluated at the mode (Laplace approximation). The column labeled “Baseline” shows the parameter estimates for the baseline model discussed in the previous section. Column “I” shows the estimates for the model using the PLM suggested by Slobodyan and Wouters (2012a) (i.e. an AR(2)). Column “II” shows the estimates for the model that considers the minimum set of state variable approach to described the PLM. Column “III” reports the estimates for the model that includes only the lagged term-spread in all PLM. Finally, Column “IV” shows the estimates when the learning ρ is restricted to zero.

In regard to parameter estimates, a few structural parameters (e.g. habit formation parameter, h, Frisch elasticity, σ_l, and steady-state markup, Φ), and all policy parameter estimates are observed to be fairly robust across the alternative specifications of the PLM studied. However, there are also some important differences. In particular, the estimates of most parameters that measure endogenous persistence (ξ_p, φ, and ψ) are somewhat sensitive to the specification used. Nevertheless, these parameter estimates are rather similar for the two PLM formulations that use the term spread (Baseline and "III"), indicating that the weakness of these endogenous sources of persistence is robust when the term spread is introduced into the PLM.

The parameter estimates of the shock processes are also observed to be quite stable across PLM formulations. However, there are a few noticeable differences. Thus, the persistence of risk premium and investment shocks and the size of their innovations are much greater for the PLM specifications that include the term spread (Baseline, "III" and "IV") than for the others. This high sensitivity of shock process estimates reinforces the suggestion discussed above that these disturbances may capture reduced-form shocks rather than truly structural shocks (Chari, Kehoe and McGrattan, 2009).

7 Conclusions

In this paper we extend the adaptive learning model of Slobodyan and Wouters (2012a) by introducing the term structure of interest rates. Our extension enables the term spread of interest rates to fully characterize the expectations of all forward-looking variables with term structure information, which is observed in real time, while retaining the features of adaptive learning based on small forecasting models and the process of updating learning coefficients based on actual values of forward-looking variables. This extension overcomes, to some extent, a general shortcoming of estimated adaptive learning versions of DSGE models based only on final revised data when in reality agents only have access to real-time data when updating their expectations over time.

Our estimation results show that including the term structure in the adaptive learning model results in decreases in the importance of a few endogenous sources of aggregate persistence (such as price and wage stickiness and the elasticity of the cost of adjusting capital). The rationale for this finding is rather simple. The extension of the estimated DSGE model introduces a multi-period forecasting environment, which means that learning dynamics are much more persistent than in one-period-ahead forecasting models (e.g. Slobodyan and Wouters, 2012a,b). The importance of multi-period forecasting is in line with results found in Eusepi and Preston (2011) for a real business cycle model. Moreover, our estimated DSGE model with term structure information shows that term premium innovations are an important source of persistent fluctuations in the nominal variables (inflation and the short-term interest rate). This finding stands in sharp contrast to the lack of transmission of term premium shocks to the macroeconomy under rational expectations.

This paper should be viewed as a first step towards understanding the importance of relying on multi-period adaptive learning based on real-time data to characterize aggregate persistence. In a follow-up paper, Aguilar and Vázquez (2018), we investigate an extended version of our model in which AL and the term structure features from the whole yield curve are combined with real-time macroeconomic information to determine agents' expectations and their decisions.

References

- Adam, Klaus (2005) Learning to forecast and cyclical behavior of output and inflation. *Macroeconomic Dynamics* 9, 1-27.

- Adam, Klaus (2007) Optimal monetary policy with imperfect common knowledge. *Journal of Monetary Economics* 54, 267-301.

- Adam, Klaus and Albert Marcet (2011) Internal rationality, imperfect market knowledge and asset prices. *Journal of Economic Theory* 146, 1224-1252.

- Aguilar, Pablo and Jesús Vázquez (2015) The role of term structure in an estimated DSGE model with learning. IRES, Université catholique de Louvain, Discussion paper 2015007.

- Aguilar, Pablo and Jesús Vázquez (2018) Term structure and real time learning. Banco de España, Working paper 1803.

- Ang, Andrew, Monica Piazzesi and Min Wei (2006) What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131, 359-403.

- Bekaert, Geert, Seonghoom Cho and Antonio Moreno (2010) New-Keynesian macroeconomics and the term structure. *Journal of Money, Credit and Banking* 42, 33-62.

- Branch, William A. and George W. Evans (2006) A simple recursive forecasting model. *Economics Letters* 91, 158-66.

- Canova, Fabio and Luca Gambetti (2010) Do expectations matter? The Great Moderation revisited. *American Economic Journal: Macroeconomics* 2, 183-205.

- Casares, Miguel and Jesús Vázquez (2016) Data revisions in the estimation of DSGE models. *Macroeconomic Dynamics* 20, 1683-1716.

- Chari, V.V., Patrick J. Kehoe and Ellen R. McGrattan (2009) New Keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics* 1, 242-266.

- Coenen, Günter, Andrew Levin and Volker Wieland (2005) Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review* 49, 975-1006.

- De Graeve, Ferre, Marina Emiris and Rafael Wouters (2009) A structural decomposition of the US yield curve. *Journal of Monetary Economics* 56, 545-559.

- Del Negro, Marco and Stefano Eusepi (2011) Fitting observed inflation expectations. *Journal of Economic Dynamics and Control* 35, 2105-2131.

- Dewachter, Hans, Leonardo Iania and Marco Lyrio (2011) A New-Keynesian model of the yield curve with learning dynamics: A Bayesian evaluation. MPRA paper 34461.

- Estrella, Arturo and Frederic S. Mishkin (1997) The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review* 41,1375–1401.

- Eusepi, Stefano and Bruce Preston (2011) Expectations, learning, and business cycle fluctuations. *American Economic Review* 101, 2844-2872.

- Evans, George W. and Seppo Honkapohja (2001) *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.

- Fama, Eugene (1990) Term structure forecast of interest rates, inflation and real returns. *Journal of Monetary Economics* 25, 59-76.

- Galí, Jordi, Frank R. Smets and Rafael Wouters (2012) Unemployment in an estimated new Keynesian model. In: Daron Acemoglu and Michael Woodford (eds.), *NBER Macroeconomics Annual*, vol. 26, pp. 329-360. Cambridge, MA: MIT Press.

- Hommes, Cars and Mei Zhu (2014) Behavioral learning equilibria. *Journal of Economic Theory* 150, 778-814.

- Hördahl, Peter, Oreste Tristani and David Vestin (2006) A joint econometric model of macroeconomic and term structure dynamics. *Journal of Econometrics* 131, 405-444.

- Ireland, Peter N. (2004) A method for taking models to the data. *Journal of Economic Dynamics and Control* 28, 1205-1226.

- Levine, Paul , Joseph Pearlman, George Perendia and Bo Yang (2012) Endogenous persistence in an estimated DSGE model under imperfect information. *Economic Journal* 122, 1287-1312.

- Mackowiak, Bartosz and Mirko Wiederholt (2009) Optimal sticky prices under rational inattention. *American Economic Review* 99, 769-803.

- Mankiw, N. Gregory and Ricardo Reis (2002) Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295-1328.

- Marcet, Albert and Thomas J. Sargent (1989) Convergence of least-squares learning in environments with hidden states variables and private information. *Journal of Political Economy* 97, 1306-1322.

- McCallum, Bennett T. (1994) Monetary policy and the term structure of interest rates. NBER Working Paper 4938. Reprinted in the *Federal Reserve Bank of Richmond Economic Quarterly*, Volume 91/4 Fall 2005.

- Milani, Fabio (2007) Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics* 54, 2065-2082.

- Milani, Fabio (2008) Learning, monetary policy rules, and macroeconomic stability. *Journal of Economic Dynamics and Control* 32, 3148-3165.

- Milani, Fabio (2011) Expectation shocks and learning as drivers of the business cycle. *Economic Journal* 121, 379-401.

- Mishkin, Frederic S. (1991) The information in the longer maturity term structure about future inflation. *Quarterly Journal of Economics* 105, 815-828.

- Ormeño, Arturo and Krisztina Molnár (2015) Using survey data of inflation expectations in the estimation of learning and rational expectations models. *Journal of Money, Credit, and Banking* 47, 673-699.

- Orphanides, Athanasios and John C. Williams (2005a) Inflation scares and forecast-based monetary policy. *Review of Economic Dynamics* 8, 498-527.

- Orphanides, Athanasios and John C. Williams (2005b) The decline of activist stabilization policy: natural rate misperceptions, learning, and expectations. *Journal of Economic Dynamics and Control* 29, 1927-1950.

- Orphanides, Athanasios and John C. Williams (2008) Learning, expectations formation, and the pitfalls of optimal control monetary policy. *Journal of Monetary Economics* 55 (Supplement), S80-S96.

- Orphanides, Athanasios and Min Wei (2012) Evolving macroeconomic perceptions and the term structure of interest rates. *Journal of Economic Dynamics and Control* 36, 239-254.

- Preston, Bruce (2005) Learning about monetary policy rules when long-horizon expectations matters. *International Journal of Central Banking* 1, 81-126.

- Pruitt, Seth (2012) Uncertainty over models and data: the rise and fall of American inflation. *Journal of Money, Credit and Banking* 44, 341-365.

- Reis, Ricardo (2009) A sticky information general equilibrium model for policy analysis. In Klaus Schmidt-Heubel and Carl Walsh (eds.), *Monetary Policy under Uncertainty and Learning*, chapter 8, pp. 227-283. Central Bank of Chile.

- Rudebusch, Glenn D. and Tao Wu (2008) A macro-finance model of the term structure, monetary policy and the economy. *Economic Journal* 118, 906-926.

- Sims, Christopher A. (2003) Implications of rational inattention. *Journal of Monetary Economics* 50, 665-690.

- Sinha, Arunima (2015) Government debt, learning and the term structure. *Journal of Economic Dynamics and Control* 53, 268-289.

- Sinha, Arunima (2016) Learning and the yield curve. *Journal of Money, Credit and Banking* 48, 513-547.

- Slobodyan, Sergey and Rafael Wouters (2008) Estimating a medium–scale DSGE model with expectations based on small forecasting models. *Mimeo.*

- Slobodyan, Sergey and Rafael Wouters (2012a) Learning in a medium-Scale DSGE model with expectations based on small forecasting models. *American Economic Journal: Macroeconomics* 4, 65-101.

- Slobodyan, Sergey and Rafael Wouters (2012b) Learning in an estimated medium-scale DSGE model. *Journal of Economic Dynamics and Control* 36, 22-46.

- Slobodyan, Sergey, and Rafael Wouters (2013) Survey expectations and learning. *Mimeo.*

- Smets, Frank R. and Rafael Wouters (2007) Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* 97, 586-606.

- Vázquez, Jesús, Ramón María-Dolores and Juan M. Londoño (2013) On the informational role of term structure in the US monetary policy rule. *Journal of Economic Dynamics and Control* 37, 1852-1871.

Part II

Term Structure and Learning

1 Introduction

This paper explores how the informational content of the term structure of interest rates transmits intro the macroeconomy through agents' expectations using a model with adaptive learning (AL) expectations.

Since the pioneering publications by Marcet and Sargent (1989) and Evans and Honkapohja (2001) a growing literature (including Preston, 2005; Milani, 2007, 2008, 2011; Eusepi and Preston, 2011; Slobodyan and Wouters, 2012a,b) has considered AL as an alternative to the rational expectations (RE) assumption in characterizing highly persistent macroeconomic dynamics. Recent papers (Sinha 2015, 2016) focus on the implications of AL in the estimation of the yield curve, but yet, there is still little evidence on how term structure information may interact with both learning and macroeconomic dynamics. This research tries to bridge this gap through a DSGE model under AL with multi-period forecasting based on the term structure information.[1]

More precisely, we consider that short-, medium- and long-term bond yields are priced with separate Euler equations. Considering the standard consumption-based asset pricing equation associated with each maturity leaves the Expectation Hypothesis of the term structure of interest rates and results into a model with multi-step forecasting (Bhansali, 2002; Jordà, 2005) to determine the multi-period-ahead expectations. This approach establishes a direct link between consumption and inflation expectations and bond yields for different

With Jesús Vázquez (UPV/EHU)

[1]This paper builds on the AL models of Slobodyan and Wouters (2012a) and Aguilar and Vázquez (2021) by considering short-, medium- and long-term bond yields priced with separate Euler equations. These two papers introduce AL into the medium-scale DSGE model suggested by Smets and Wouters (2007) (hereinafter called the SW model), but Aguilar and Vázquez (2021) extend the Slobodyan and Wouters (2012a) framework by both incorporating the short-end of the term structure of interest rates and imposing the expectations hypothesis.

yield-horizons and it allows us to exploit the informational content of the term structure by assuming that agents' expectations convey information onthe associated observed yield. This approach has the advantage of allowing us to define simple and parsimonious learning models. The use of term structure information in agents' forecasting models is further motivated by the ability of term spreads to forecast inflation (Mishkin, 1990) and real economic activity (Estrella and Hardouvelis, 1991, Estrella and Mishkin, 1997).

Our AL approach falls under the broad class of a restricted perceptions equilibrium, where agents use a small misspecified model but form their beliefs optimally given the misspecification (Sargent, 1991; Hommes and Sorger, 1998; Honkapohja, Mitra, and Evans, 2012), however, using multi-step forecasting deviates from the two main approaches in the recent literature to AL. One approach (called the Euler equation approach) focuses on short-sighted agents, for whom optimal current decisions are based on just one-period-ahead expectations showing up in the standard Euler equations (e.g. Milani, 2007; Slobodyan and Wouters, 2012a,b), while the other approach focuses on long-sighted agents (e.g. Preston, 2005; Eusepi and Preston, 2011; Sinha, 2015; and Sinha; 2016), as under RE, taking into account infinite-horizon forecasts driven by their intertemporal decision problem, that amplifies the source of persistence dynamics (see Eusepi and Preston (2011)). By including the long-end of the term structure of interest rates, our approach certainly goes beyond the one-period-ahead expectations, but still follows the Euler equation approach. Additionally, direct multi-step forecasting has the advantage of introducing additional flexibility, resulting into a better model fit, which is ignored when considering iterated forecast when dealing with misspecified models, where misspecification errors in the beliefs are compounded with the forecast horizon, as discussed in Jordà (2005).

In spite of the highly restricted information set assumed in the forecasting models used by agents, as they only observe bond yields to form their expectations, our estimated AL-DSGE model extended with term structure information fits the data much better than the RE version of the model or than other AL specifications based on larger information set as in Slobodyan and Wouters (2012a). In particular, it provides a good fit of macroeconomic dynamics, the yield curve and, the model fit measured by the log data density. The rationale behind this result is the additional persistence introduced from adding multi-period Euler equations associated with the different bond yields under AL that enhances the expectation channel in the model, standing in sharp contrast to RE, where multi-period Euler equations are redundant for agents expectations. Our modeling approach results in an estimated 10-year AL term premium showing a contemporaneous correlation of roughly 0.9 with the estimated term premium measure of Adrian, Crump and Moench (2013) obtained from a five-factor no-arbitrage affine term structure model. This finding is quite remarkable since DSGE models under RE typically need both a third-order Taylor approximation and Epstein and Zin preferences (Epstein and Zin, 1989) featuring a huge risk aversion to reproduce a sound term premium (Rudebusch and Swanson, 2012) whereas our AL term premium based on a first-order approximation is obtained assuming logarithmic preferences.[2]

The rest of the paper is structured as follows. Section 2 briefly describes the DSGE model with AL based on term structure information under the multi-period forecasting hypothesis. Section 3 shows the estimation results and discusses their implications. Section 4 analyzes the term premium dynamics implied by the AL and the RE versions of the model compared to non-arbitrage affine models of the term structure . Section 5 shows that our results are robust when expectations are disciplined with SPF and how our model derives expectations with similar properties to those observed in survey data. Section 6 concludes.

2 An AL model with term structure

We include multi-period forecasting in a model with AL to investigate the interaction between the yield curve—an important source of publicly available information observed — and the

[2]There is a caveat though: Our estimated term premium measure is purely exogenous, as in De Graeve, Emiris, and Wouters (2009). In a follow-up paper, Vázquez (2019) focuses on a standard measure of the term premium defined as the excess return that a consumer earns in expectation by buying a long-term bond and holding it to maturity instead of buying short-term bonds and rolling them over until the maturity date of the long-term bond.

agents' learning process in an estimated DSGE model. Our model builds on the SW model and its AL extensions studied by Slobodyan and Wouters (2012a) and Aguilar and Vázquez (2021). This standard medium-scale estimated DSGE model contains both nominal and real frictions affecting the choices of households and firms. To exploit the informational content of the yield curve we price bond yields using the standard consumption-based asset pricing equation associated with each maturity and use them to form the associated expectation appearing in that equation. Considering the standard consumption-based asset pricing equation associated with each maturity leaves the Expectation Hypothesis of the term structure of interest rates from the previous chapter and explicitly turns into multiple-period-ahead expectations. This enhances the channel of expectations in the model as pricing with the Euler equation relates the term premium to a wedge between the optimal yield and the expectations about future consumption and inflation. This differs from the approach followed in a parallel paper—Vázquez and Aguilar (2021)—which imposes the term structure expectation hypothesis obtained by iterating the asset pricing equations forward. Our approach enables us to abandon the term structure expectation hypothesis while still considering the possibility that multiple-period-ahead expectations under AL may matter to agents' decisions (up to the 10-year yield).[3]

The term structure extension

Formally, the first-order conditions describing the optimal decisions of the representative consumer in the SW model imply the standard consumption-based asset pricing equation

[3] As pointed out by Eusepi and Preston (2018, footnote 10), expected yields (returns) under arbitrary subjective beliefs may not satisfy no-arbitrage with multiple bond maturities (assets). They impose the expectations hypothesis of the term structure to overcome this issue as we did in our previous paper (Aguilar and Vázquez (2021)). Although this might be a cause for concern, we focus on the (optimal) yields implied by the separate Euler equations in this paper. The main reason for this is that term premia on medium- and long-term nominal bonds are viewed as compensation for inflation and consumption risks faced by investors over the lifespan of bonds—Rudebusch and Swanson (2012). Hence, the separate Euler equation approach precisely highlights this connection between the term premium and the nominal and real risks (see equation (1) below) by relating the term premium to a wedge between the optimal yield and the expectations about future inflation and consumption profiles. In contrast, the expectations hypothesis of the term structure links the term premium to a different wedge defined as the difference between a long-term yield and the expected average of the short-term rate over the lifespan of the long-term yield.

associated with each maturity:

$$E_t\left[\beta^j \frac{U_C(C_{t+j}, N_{t+j}) exp(\xi_t^{\{j\}})(1+R_t^{\{j\}})^j}{U_C(C_t, N_t)\prod_{k=1}^{j}(1+\pi_{t+k})}\right] = 1,\ for\, j =\ 1,\ 2,\ ...,\ n,$$

where E_t denotes the RE or the AL operator depending on the estimated model analyzed below, β is the discount factor, U_C denotes the marginal utility of consumption, and C_t, N_t, $R_t^{\{j\}}$, and π_t stand for consumption, labor, the nominal yield of a j-period maturity bond, and the inflation rate, respectively.[4]

Assuming that the utility function is logarithmic in consumption, the (linearized) consumption-based asset pricing equations can be written as

$$\left(\frac{1}{1-\frac{h}{\bar{\gamma}}}\right)c_t - \left(\frac{\frac{h}{\bar{\gamma}}}{1-\frac{h}{\bar{\gamma}}}\right)c_{t-1} =$$

$$E_t\left[\left(\frac{1}{1-\frac{h}{\bar{\gamma}}}\right)c_{t+j} - \left(\frac{\frac{h}{\bar{\gamma}}}{1-\frac{h}{\bar{\gamma}}}\right)c_{t+j-1}\right] - \left[j r_t^{\{j\}} - E_t\sum_{k=1}^{j}\pi_{t+k} + \xi_t^{\{j\}}\right],\ for j = 1, 2, ..., n, \quad (1)$$

where lower case variables represent log-deviations from balanced-growth values or, alternatively, deviations from steady-state values. In particular, for $j = 1$ the last expression results in the standard IS-curve equation in the SW model assuming a logarithmic utility function in consumption—i.e. Equation (2) in Smets and Wouters (2007, p.588) when the risk aversion parameter, σ_c, is one.[5]

[4]The inclusion of a risk premium shock for each maturity, $\xi_t^{\{j\}}$, is in line with the view of many authors of interpreting the wedge between the model-implied yield, $R_t^{\{j\}}$, and the observed yield as a measure of fluctuations in the risk premium (e.g. De Graeve, Emiris and Wouters, 2009). Moreover, since we focus on government bonds in our empirical analysis, $\xi_t^{\{j\}}$ can also be understood as a *convenience yield term* (see, among others, Krishnamurthy and Vissing-Jorgensen, 2012; Greenwood et al., 2015; Del Negro et al., 2017) defined as a premium related to the safety and liquidity attributes of (US) government bonds relative to assets with the same payoff, but without such singular features.

[5]DSGE models under RE typically require a huge value of the risk aversion parameter close to 100 to reproduce a sizable *endogenous* term premium based on a third-order approximation of the model (Rudebusch and Swanson, 2012). By assuming $\sigma_c = 1$, we want to stress the potential of AL to generate a sound term premium.

Subtracting the expression in (1) for $j = 1$ from (1), we obtain the following expression linking the term spread defined by the j- and the 1-period yields with consumption growth and inflation expectations:

$$r_t^{\{j\}} - r_t^{\{1\}} = \left(\frac{j-1}{j}\right) r_t^{\{1\}} + \frac{1}{j} E_t \left[c \left(c_{t+j} - c_{t+1} \right) + (1-c) \left(c_{t+j-1} - c_t \right) \right]$$

$$+\frac{1}{j} E_t \sum_{k=2}^{j} \pi_{t+k} - \frac{1}{j} \left(\xi_t^{\{j\}} - \xi_t^{\{1\}} \right), \tag{2}$$

where $c = \frac{1}{1-\frac{h}{\bar{\gamma}}}$. This optimality condition clearly shows that term spreads must be linked to consumption and inflation expectations in equilibrium, which rationalizes our modeling approach of using term structure information to characterize the formation of agents' expectations, which motivates the small forecasting model considered below. In particular, it links the j-period term spread with the forecasts of consumption growth and inflation up to the j-period horizon. As a consequence, the term premium can be defined as the wedge between the optimal yield and the expectations about future consumption and inflation.

Under RE, the optimality conditions in (1) for $j > 1$ are somewhat redundant because long-term bonds are redundant assets in equilibrium. However, agents face much greater uncertainty under AL than under RE (i.e. AL agents know neither the true model nor the alternative sources of uncertainty, which are assumed to be known under RE). Consequently, long term bonds can be viewed as more useful to hedge against future sources of uncertainty under AL than under RE. Moreover, as is clear from the set of consumption-based asset pricing equations (1), optimal consumption household decisions explicitly involve the multi-period forecasting hypothesis since current consumption depends, among other things, on the expected paths of both future consumption and inflation.[6]

[6]The approach followed in this paper of yield pricing with separate Euler equations stands in sharp contrast to the approach followed in Aguilar and Vázquez (2021) and Vázquez and Aguilar (2021) where the expectations hypothesis of the term structure is imposed. The former imposes a tight connection between long-term yields and the expected paths of future consumption and inflation, whereas the latter imposes a close connection between the long term-yield and the expectations for the short-term rate (i.e. the expectations about future monetary policy). Of course, when there are no risk premium shocks the two approaches are tantamount under RE, but they need not be equivalent under AL.

As in the SW model, we assume that the risk premium shock $\xi_t^{\{1\}}$ follows an AR(1) process: $\xi_t^{\{1\}} = \rho^{\{1\}}\xi_{t-1}^{\{1\}} + \eta_t^{\{1\}}$, but the remaining risk premium shocks $\xi_t^{\{j\}}$, for $j > 1$, are assumed to follow AR(1) processes augmented with an additional component that allows for an interaction with the short-term risk premium shock: $\xi_t^{\{j\}} = \rho^{\{j\}}\xi_{t-1}^{\{j\}} + \rho_\xi^{\{j\}}\eta_t^{\{1\}} + \eta_t^{\{j\}}$, where $\rho_\xi^{\{j\}}$ captures the interaction of the short-term risk premium innovation, $\eta_t^{\{1\}}$, with the term premium shock, $\xi_t^{\{j\}}$, associated with the j-period maturity bond. In this simple way, term premium shocks are correlated across alternative maturities, which helps to explain the imperfect but large correlation between yields associated with alternative maturities.

We define the risk-free yield, $r_t^{f\{j\}}$, associated with the j-period maturity as that implied by (1) with $\xi_t^{\{j\}} = 0$. Hence, a term premium measure, $tp_t^{\{j\}}$, associated with the j-period maturity bond is given by

$$tp_t^{\{j\}} \equiv r_t^{\{j\}} - r_t^{f\{j\}} = -\frac{1}{j}\xi_t^{\{j\}},$$

where both yields and the term premium are measured in deviations from their respective steady-state values. This definition of the term premium is purely exogenous, and is similar to the one considered in De Graeve, Emiris and Wouters (2009) based on measurement errors.[7]

Forming expectations

The AL literature assumes that agents do not know exactly the full structure of the model and they need to estimate a forecasting model using a limited set of information to form their expectations about the behavior of the economy. This type of learning enable us to address the informational content of TS, as we assume that it is their only information available.[8] Traditionally, there are two alternatives to defining the forecasting model, one assumes that

[7] Nonetheless, alternative definitions of the bond term premium may be worth considering. In a follow-up paper, Vázquez (2019) studies a standard definition of the term premium defined by the wedge between the optimal yield of a long-term bond and the return obtained by buying short-term bonds and rolling them over until the maturity date of the long-term bond—the latter being that implied by the expectation hypothesis of the term structure.

[8] AL lies within the bounded rationality literature, as agents use models that are misspecified and the choice of their forecasting rules is not a part of the optimal decision-making process of the agents, as oppossed to the internal rationality models (see Adam and Marcet (2011)). For the purpose of this paper deciding on the PLM allow us to explicitly capture the potential of TS as a source of information into agents expectations.

agents' Perceive Law of Motion (PLM) has the same functional form of the RE law of motion, i.e. agents observe the state variables, but the true coefficients are unknown for them, as in Evans and Honkapohja (2001), in Preston (2005), Milani (2007, 2008) and Eusepi and Preston (2011). The second alternative assumes that agents use small forecasting models, i.e. misspecified models, whose observed variables are the relevant macroeconomic variables in the model, as in, Milani (2011), Slobodyan and Wouters (2012a,b and 2017) and Aguilar and Vázquez (2021). In our case, to address the informational content of the term structure, we assume that agents' forecasting models only include term spreads. As emphasized above, a PLM based on term structure information is motivated by the interaction between term spreads and the expectations for both consumption and inflation as shown in equation (2). Moreover, term spreads help to predict inflation (Mishkin, 1990) and real economic activity (Estrella and Hardouvelis, 1991, Estrella and Mishkin, 1997).[9]

A PLM with only term structure information

As emphasized above, a PLM based on term structure information is motivated by the interaction between term spreads and the expectations for both consumption and inflation as shown in equation (2).

Formally, we use the 2-quarter and 4-quarter term spreads to characterize the PLM of all forward-looking variables up to the 4-quarter horizon as follows:

$$\begin{cases} E_t y_{t+1} = \theta_{y,t-1} + \beta_{y,t-1}^{\{2\}} sp_{t-1}^{\{2\}}, \quad for\ \ y = i,\ r^k,\ q,\ w \\ E_t y_{t+j} = \theta_{y,t-1}^{\{j\}} + \beta_{y,t-1}^{\{j,2\}} sp_{t-1}^{\{2\}}, \quad for\, y =\ c,\ \pi\ and\, j =\ 1,\ 2,\ 3 \\ E_t y_{t+j} = \theta_{y,t-1}^{\{j\}} + \beta_{y,t-1}^{\{j,4\}} sp_{t-1}^{\{4\}}, \quad for\, y =\ c,\ \pi\ and\, j = 4 \end{cases} \tag{3}$$

[9] McCallum (1994) also exclusively uses term spreads as simple predictors for future macroeconomic conditions for defining monetary policy rules.

where i, r^k, q, w, c and π stand for (in deviation from their respective steady-state values or detrended by the balanced growth rate) investment, rental rate of capital, Tobin's q, real wage, consumption, and inflation, respectively; and $sp_{t-1}^{\{2\}} = r_{t-1}^{\{2\}} - r_{t-1}^{\{1\}}$ and $sp_{t-1}^{\{4\}} = r_{t-1}^{\{4\}} - r_{t-1}^{\{1\}}$ denote the term spreads associated with the 2-quarter and 1-year term spreads, respectively, measured with respect to the 1-quarter interest rate.[10]

The PLM follows a simple rule of thumb of determining the short-term expectations up to 3-quarters ahead with the short-term term spread (i.e. $sp_{t-1}^{\{2\}}$) and the longer-term expectations with the 1-year term spread ($sp_{t-1}^{\{4\}}$). This is somewhat in line with the optimality condition (2), which establishes that the j-period term spread, $sp_t^{\{j\}}$, is linked to j-periods-ahead expectations for consumption and inflation. Moreover, the presence of time-varying intercepts $\theta_{y,t-1}^{\{j\}}$ relaxes the RE assumption of agents having perfect knowledge about a common deterministic growth rate and a constant inflation target made in the Smets and Wouters (2007) model—i.e. a time-varying intercept coefficient allows expectations to trace low-frequency shifts in the data.

In spite of the simplicity of the PLM, the findings shown below provide robust support for our estimated AL model based only on term structure information since (i) it gives a better model fit than the RE and other AL versions; (ii) it does a reasonable job in reproducing business cycle features; and (iii) the implied AL term premium is closely correlated with the term premium estimates obtained from reduced-form models such as no-arbitrage affine term structure models.

Long-term expectations

Considering a long term maturity yield such as the 10-year yield means characterizing expectations of consumption and inflation up to a 40-quarter horizon. In the current setup a long forecasting horizon dramatically increases the number of expectation functions of consump-

[10]In the estimation exercise below, we consider that the federal funds interest rate is equivalent to the 1-quarter interest rate. This assumption is standard in the related literature since these two interest rates are closely correlated.

tion and inflation, potentially leading to a curse of dimensionality problem. To deal with this issue we assume the following simple recursive structure for consumption and inflation expectations on forecast horizons beyond the four-quarter horizon:

$$\begin{cases} E_t c_{t+j} = \mu_c E_t c_{t+j-1}, \\ E_t \pi_{t+j} = \mu_\pi E_t \pi_{t+j-1}, \end{cases}$$

for $j > 4$. This structure builds on the forecasting rules described above which, among others, characterize $E_t c_{t+4}$ and $E_t \pi_{t+4}$. The parameters μ_c and μ_π are estimated jointly with the rest of model parameters.

As in the Euler equation learning approach, the PLM for each horizon is estimated separately according to the forecasting models assumed and thus they do not have to be consistent with each other up to the 4-period (quarter) horizon. Our AL approach considers direct multi-step forecasting (Bhansali, 2002; Jordà, 2005) to determine the expectations up to the 4-quarter horizon. This is in contrast to the approach followed in the related AL literature of using iterated forecasts, which are built in general on a misspecified model. As discussed in Jordà (2005), among others, direct forecasts associated with long-term horizons outperform iterated forecasts when dealing with misspecified models since misspecification errors are compounded with the forecast horizon. Direct multi-step forecasting introduces additional flexibility, which is ignored by considering iterated forecasts, and this extra flexibility may result in a better model fit.

The flexibility of the PLM used in this paper across forecast horizons enables us to overcome the log-linear approximation typically imposed in DSGE modeling to some extent. Thus, this flexibility improves model fit by capturing non-linear features as well as low-frequency patterns in the data (e.g. the downtrend of inflation in the last thirty years).[11]

[11] Notwithstanding the flexibility of having agents forecasting the real yield associated with each maturity by using a different forecasting model, one might ask what ensures that there is a stationary solution to this extended model. This is not much of an issue since unbounded solutions are ruled out in the implementation of the Kalman filter learning algorithm used in this paper. Technically, the reason is that the so called *projection facility* used in the learning algorithm prevents changing beliefs from moving the system dynamics

We start by restricting our attention to the four-period (quarter) forecasting horizon. By doing this we certainly move beyond the 1-period forecast horizon considered in many related AL models. Later, we move to investigate the case of finite-long-sighted consumers by using a 10-year forecasting horizon in their decision making along with four additional medium- and long-term yields (up to the 10-year Treasury yield) as observables in order to discipline long-term expectations. In this way, we can assess how much aggregate persistence can be generated with a multi-period forecasting model based only on short-sighted consumers versus a model based on long-sighted consumers.

3 Estimation results

We begin this section by describing the data and the estimation approach, then proceed to discuss the model fit, estimation results, the variance decomposition of shocks, the PLM of the AL, and the implied expectations for inflation and the consumption growth rate.

3.1 Data and estimation approach

Our AL model extended with term structure and the associated RE model are estimated using US data for the great moderation period running from 1983:4 until 2008:1.[12] The set of observable variables is the same as the one used by Slobodyan and Wouters (2012a) (i.e. the quarterly series of the inflation rate, the Fed funds rate, the log of hours worked, and the quarterly log differences of real consumption, real investment, real wages, and real GDP) with the addition of the 1-year zero-coupon Treasury yield. Later, we extend the set of observables to include first four additional yields up to the 10-year (zero-coupon Treasury bond) yield. The zero-coupon Treasury bond yields come from the Gürkaynak et al. (2007)

into an unstable region.

[12] Our estimated model considers consumption growth forecasts from the SPF in addition to the inflation forecasts used by Ormeño and Molnár (2015). While inflation forecasts are reported back to the late 1960's, the consumption growth forecast time series starts at 1981:3. We decided to start our sample at 1983:4. In this way, we ignore the inflationary episode right after the Volcker monetary experiment and focus on the great moderation period when the policy rule was well characterized by a Taylor-type rule.

dataset available on the Board of Governors of the Federal Reserve's research data website.

The measurement equation is

$$X_t = \begin{bmatrix} dlGDP_t \\ dlCONS_t \\ dlINV_t \\ dlWAG_t \\ dlP_t \\ lHours_t \\ FEDFUNDS_t \\ k - year\,TB\,yield_t \end{bmatrix} = \begin{bmatrix} \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\pi} \\ \bar{l} \\ \overline{r} \\ \overline{r}^{\{4k\}} \end{bmatrix} + \begin{bmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - i_{t-1} \\ w_t - w_{t-1} \\ \pi_t \\ l_t \\ r_t \\ r_t^{\{4k\}} \end{bmatrix}, \tag{4}$$

where l and dl represent the log and the log difference, respectively. $\overline{\gamma} = 100(\gamma - 1)$ is the common quarterly trend growth rate for real GDP, real consumption, real investment, and real wages. $\bar{l}$, $\overline{\pi}$, $\overline{r}$ and $\overline{r}^{\{4k\}}$ are the steady-state levels of hours worked, inflation, the federal funds rate, and the k-year ($4k$-quarter) bond yield, respectively.

We follow a Bayesian estimation procedure. First, the log posterior function is maximized by combining prior information on the parameters with the likelihood of the data. The prior assumptions are exactly the same as in Slobodyan and Wouters (2012a). The Metropolis-Hastings algorithm is used to generate the posterior distribution and to compute the log density of the model.[13]

3.2 Posterior estimates

Our estimated AL with term structure—henceforth, ALTS—model differs from the Slobodyan and Wouters (2012a)—henceforth, SlW—model in the incorporation of the term structure and its use to form expectations, as highlighted at the beginning of Section 2. In

[13]The DSGE models are estimated using Dynare codes kindly provided by Sergey Slobodyan and Raf Wouters with a set of modifications to accommodate the presence of term structure information in both the structural model and the small forecasting models, as described above.

order to facilitate the discussion, we start by focusing on the short-end of the term structure and subsequently extend the analysis to consider the yield curve up to the 10-year yield.

Short-end of the term structure

Table 1 shows the estimation results for the alternative samples and specifications considered. The first column of Table 1 shows the estimation results of the SlW model for the sample period 1984-2007, column 2 shows the estimate of the SlW model including the 1-year Treasury bill as an observable—i.e. using 8 time series as observables. The corresponding estimates are reported in column 2. Column 3 contains the estimation results for our ALTS model using these 8 time series as observables.

For each model estimated, Table 1 firstly reports the number of observable time series, and the model fit based on marginal likelihood. The remaining rows show the posterior mean and the corresponding 90 percent interval of the posterior distribution—in parentheses—for four groups of selected parameters. The first and second groups include the parameters featuring real and nominal rigidities, respectively. The third group includes the parameters which describe the moving average coefficients governing price and wage markup shocks. Finally, the fourth group includes the key policy rule parameters.

A comparison of columns 1 and 2 shows that including the 1-year Treasury bill as an observable in the SlW model decreases the importance of all sources of endogenous rigidity. The rationale for this generalized decrease is that by considering multi-period expectations as in (1), current consumption is linked to longer expectation paths of both inflation and consumption, which induces greater endogenous persistence and hence takes over from other sources of persistence. Moreover, there is a large increase in persistence driven by the overall increase in ARMA coefficients featuring price and wage markup shocks.

A comparison of the marginal likelihood values in columns 2 and 3 shows that the switch from the learning scheme, based on small forecasting models in SlW, to our, TS only based information -ALTS- , results in a large improvement in model fit [-427.48-(-501.41)=73.93]. Likewise, the AL model (ALTS) outperforms the model fit with respect to RE [-427.48-(-

477.45)=49.47].

Regarding the posterior estimates of parameters, most sources of endogenous persistence can be seen in general to lose a great deal of importance under AL based on term structure information. Thus, the estimates of the habit formation parameter, h, and the elasticity of the cost of adjusting capital, φ, are much smaller under AL (0.13 and 0.72, respectively) than under RE (0.89 and 5.85, respectively). Similarly, the price Calvo probability and the elasticity of capital utilization adjusting cost, ψ, are smaller under AL than under RE, whereas the opposite occurs for the price indexation parameter, ι_p, and the wage Calvo probability, ξ_w. Regarding exogenous sources of price and wage markup persistence, we find that price and wage markup shock persistence are lower in their moving average component under AL than under RE.

The main conclusion is that AL under the multi-period forecasting hypothesis substitutes, and increases, the sources explaining aggregate persistence compared to RE (and AL as in Slobodyan and Wouters (2012a), which leads to an improve in the log data density. The rationale behind this result is that introducing multi-period Euler equations associated with the different bond yields under AL enhances the expectation channel in the model, which translates into higher persistence in the model dynamics and additional flexibility to fit the data, as discussed in Eusepi and Preston (2011).[14] This stands in sharp contrast to RE, where multi-period forecasting is redundant for agents expectations.

Interestingly, this characterization results into a better fitting of the term premium, as shown in Section 4. Overall, our results further assess the findings in Milani (2007) using a small-scale DSGE model and Slobodyan and Wouters (2012a) using a medium-scale DSGE model, since these two papers only consider one-period-ahead expectations in shaping current decisions of agents and disregard the contribution of forecasts over longer horizons. In spite of the important differences in modeling learning, our estimation results are also in line with

[14]Eusepi and Preston (2011) show that taking into account infinite-horizon forecasts driven by their intertemporal decision problem amplifies the source of persistence dynamics, while at the same time direct multi-step forecasting has the advantage of introducing additional flexibility, resulting into a better model fit as discussed in Jordà (2005).

those found in Eusepi and Preston (2011) in the sense that AL with multi-period forecasting induces a much stronger mechanism for explaining aggregate persistence.

Table 1. Selected parameter estimates

	SlW (1)	SlW-TS (2)	ALTS (3)	RETS (4)
Number of observables	7	8	8	8
log data density	-424.4	-501.41	-427.48	-477.45
Parameters associated with real rigidities				
habit formation	0.82	0.47	0.13	0.89
(h)	(0.79,0.86)	(0.45,0.50)	(0.11,0.15)	(0.87,0.93)
cost of adjusting capital	7.02	1.87	0.72	5.85
(φ)	(5.62,8.32)	(1.77,1.98)	(0.61,0.86)	(4.51,6.72)
capital utilization adjusting cost	0.61	0.31	0.17	0.63
(ψ)	(0.42,0.80)	(0.27,0.34)	(0.09,0.25)	(0.44,0.82)
Parameters associated with nominal rigidities				
price Calvo probability	0.79	0.48	0.67	0.74
(ξ_p)	(0.75,0.83)	(0.44,0.51)	(0.61,0.72)	(0.71,0.77)
wage Calvo probability	0.72	0.24	0.59	0.40
(ξ_w)	(0.62,0.79)	(0.22,0.26)	(0.53,0.64)	(0.38,0.42)
price indexation	0.36	0.34	0.36	0.30
(ι_p)	(0.17,0.57)	(0.32,0.36)	(0.28,0.42)	(0.24,0.38)
wage indexation	0.29	0.36	0.35	0.36
(ι_w)	(0.09,0.48)	(0.29,0.43)	(0.29,0.42)	(0.29,0.43)
Parameters associated with price and wage markups				
markup price MA coef.	0.53	0.62	0.66	0.86
(μ_p)	(0.42,0.64)	(0.55,0.70)	(0.57,0.74)	(0.74,0.93)
markup wage MA coef.	0.54	0.19	0.19	0.64
(μ_w)	(0.38,0.70)	(0.14,0.23)	(0.12,0.27)	(0.48,0.80)
Policy rule parameters				
inertia	0.87	0.91	0.76	0.83
(ρ_r)	(0.83,0.91)	(0.90,0.92)	(0.71,0.82)	(0.82,0.85)
inflation response	1.57	1.35	1.67	1.55
(r_π)	(1.20,1.94)	(1.30,1.41)	(1.54,1.85)	(1.19,1.92)

Notes: parameter notation and standard deviation in parentheses

Long maturity term structure

Next we extend the AL and RE models to incorporate information from longer maturity term structure. More precisely, we consider 3-year, 5-year, 7-year, and 10-year zero-coupon Treasury bond yields as additional observables in the measurement equation (7) (i.e. we consider 12 observables). Moreover, four additional consumption-Euler equations are considered in the AL and RE models estimated, each associated with one of these additional yields. For instance, as implied by equation (1), the equation for the 10-year yield is given by

$$\left(\frac{1}{1-\frac{h}{\bar{\gamma}}}\right) c_t - \left(\frac{\frac{h}{\bar{\gamma}}}{1-\frac{h}{\bar{\gamma}}}\right) c_{t-1} = E_t\left[\left(\frac{1}{1-\frac{h}{\bar{\gamma}}}\right) c_{t+40} - \left(\frac{\frac{h}{\bar{\gamma}}}{1-\frac{h}{\bar{\gamma}}}\right) c_{t+39}\right]$$
$$- \left[40 r_t^{\{40\}} - E_t \sum_{k=1}^{40} \pi_{t+k} + \xi_t^{\{40\}}\right].$$

In order to facilitate comparison, Table 2 shows the log data density fit of short- and long-maturity model with AL and RE while the estimation results for the selected parameters are shown in Table 2B in the appendix. Table 2 shows that the AL model characterizes the term structure much better than the RE model. The log data density improvement obtained by using term structure information up to the 10-year yield as an observable is roughly three times greater in the AL model (325.35) than in the RE model (99.66). Compared to the short-maturity, this gain is more than four times higher. This clearly shows the impact of multi-period forecasting as discussed previously when fitting the yield curve at longer horizons. In the following section we show how this results into a greater fit of macroeconomic variables and most notably into the estimation of a 10-year AL term premium showing a high correlation with term premium measures derived from affine models.

Table 2. Model fit comparison: log data density

	ALTS-1Y	RE-1Y	ALTS-10Y	RE-10Y
Observables	8	8	12	12
Log data density	-427.48	-477.45	325.35	99.76
Model fit gain	49.97		225.59	

Focusing on the AL estimates, Table 2B in the appendix shows that the estimates of parameters featuring both endogenous and exogenous persistence are not very sensitive to considering the whole yield curve. The only noteworthy exceptions are the increases in the estimates of the elasticity of capital utilization adjusting cost, ψ, and wage indexation, ι_w. In sum, our empirical results suggests that even the model based on the short-end of the term structure is able to induce a great deal of aggregate persistence through the learning mechanism which is similar to that generated by the model based on the yield curve up to the 10-year yield, but greater than the model with only 1-quarter-ahead forecast horizons as considered in Milani (2007) and Slobodyan and Wouters (2012a) using small forecasting models.

3.3 Model fit

Along with the overall model fit based on the posterior log data density, we also analyze the performance of our estimated baseline AL model in reproducing selected second-moment statistics obtained from actual data as shown in Table 3. We focus on three types of moment: standard deviations, contemporaneous correlations with inflation, and first-order autocorrelations. Regarding the actual size of fluctuations, we observe that the AL model provides a reasonably good match for the standard deviation of inflation and the growth rates of real variables: output, consumption, investment, and wages. A similar conclusion can be drawn by looking at the correlation of inflation with these real variables. Thus, the AL model is able to reproduce negative and low correlations of inflation with real variables relatively well, apart from the correlation of real wage growth with inflation which is positive in the model. The AL model does also a good job in reproducing the actual persistence, except in the case of real wages and inflation which are higher in the model.

Table 3. Actual and simulated second moments

Actual data	Δc	Δinv	Δw	Δy	π
Standard deviation	0.50	1.54	0.70	0.56	0.23
Correlation with π	-0.22	-0.16	-0.22	-0.12	1
Autocorrelation	0.21	0.51	0.06	0.29	0.84
Simulated data	Δc	Δinv	Δw	Δy	π
Standard deviation	0.50	1.39	0.76	0.64	0.22
Correlation with π	-0.35	-0.35	0.32	-0.41	1.0
Autocorrelation	0.18	0.56	0.49	0.25	0.99

3.4 Analysis of the PLM

Figure 1 shows the trend over time of the PLM coefficients of the four-quarter-ahead expectations for inflation and consumption. We focus on the four-quarter expectation horizon since the 1-year term spread is the only observable term spread considered in the PLM. The time-varying intercept of the PLM of inflation (consumption) shows how agents' perception about steady-state inflation (balanced-growth consumption) changes over time. Thus, the intercept of inflation expectations captures the fall in inflation expectations on the long-run over the sample period, whereas the (roughly constant) intercept of consumption captures the fairly constant balanced-growth consumption expectations. The term-spread coefficients associated with these two PLM are positive, indicating that a higher 4-quarter bond yield today (relative to the 1-quarter interest rate) anticipates higher 4-quarter-ahead expectations of both inflation and consumption. In the case of consumption this coefficient is rather stable, which suggest that yield information has a steady transmission over time, whereas they tend to decrease over time in the case of inflation. The size of the term-spread consumption coefficient relative to the inflation coefficient, roughly eight times larger, shows how term structure information has an important effect on the real side of the macroeconomy.[15]

The forecasting model (i.e. the PLMs) assumed in this paper are rather simplistic, but the bottom line, as shown in the graph below, is that the time-varying intercepts of the PLMs capture the low-frequency movements of macroeconomic data (e.g. the inflation downtrend since the early 80's) well whereas the time-varying coefficients associated with term spreads seem to capture high-frequency movements of the data (growth rates of wages and investment, see in the appendix) relatively well[16] Additionally, in Section 5 we compare the model implied expectations using these PLM have with survey data, and, despite the coefficients are rather stable, the AL expectations obtained from the model show a similar behavior (in terms of volatility and trend) to those reported surveys.

[15]The estimated learning coefficients of the PLM are fairly stable. This finding is consistent with the estimated posterior confidence interval of the learning coefficient, $\rho \in [0.67, 0.75]$. As shown in Slobodyan and Wouters (2012a), in their framework a low ρ is typically associated with reduced, but sufficient, variation in beliefs. We have investigated the implications of $\rho < 0.2$, no time variation of the PLM. This restriction results into a deterioration of term premium estimation.

[16]The crucial role played by time-varying intercepts in characterizing AL dynamics has been also stressed in the related literature-for instance, Eusepi and Preston (2011, 2018) and Slobodyan and Wouters (2012a). Thus, information about the economy stemming from the yields is transmitted successfully into agents expectations.

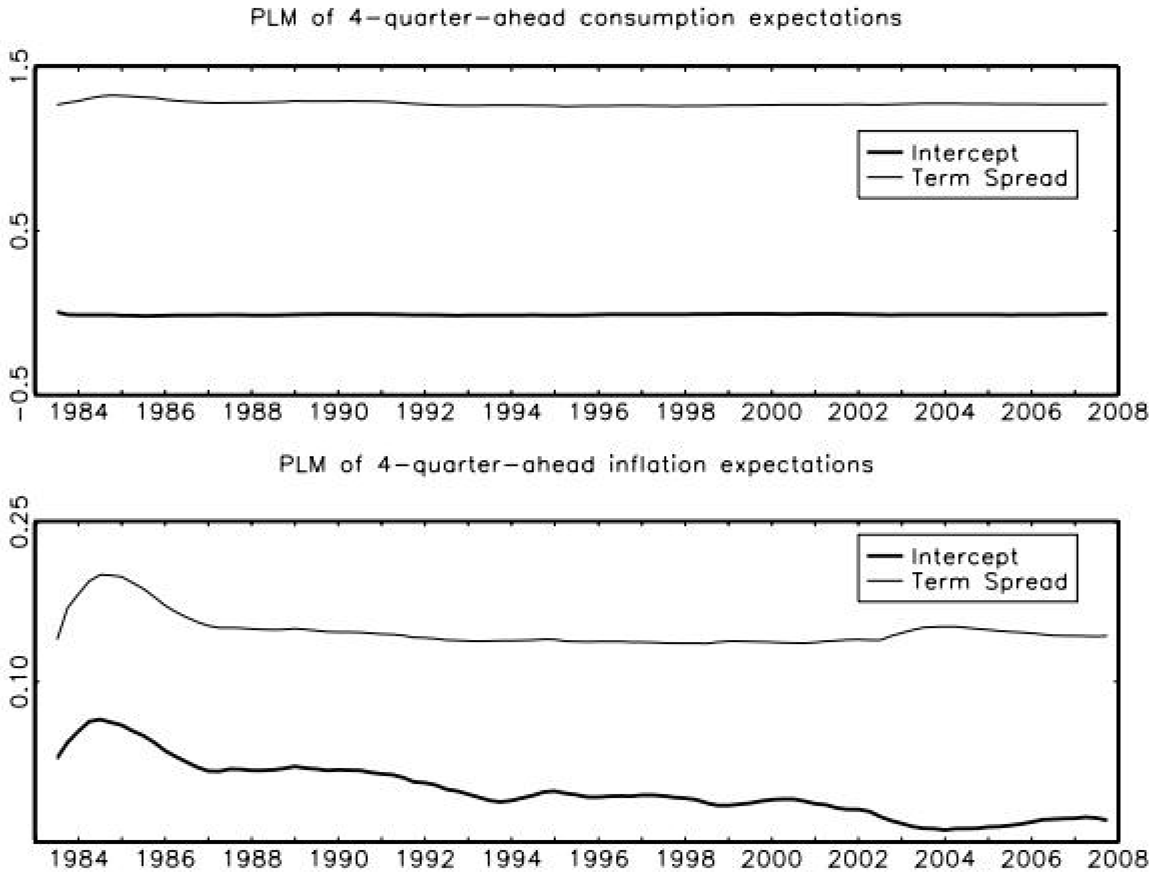


Figure 1. PLM of inflation and consumption expectations

4 Term premium analysis

This section compares the estimated AL term premia with the premium estimated from the non-arbitrage affine model of Adrian, Crump and Moench (2013)—henceforth ACM and the term premium implied by the RE model. As is standard in the literature, we focus on the term premia associated with the 10-year yield.[17]

Figure 2 shows the AL term premium measure obtained from the estimated baseline

[17]The AL term premium is given by $\left[\left(\overline{r}^{\{40\}}-\overline{r}\right)-\frac{1}{40}\xi_t^{\{40\}}\right]\times 4$ when written in annualized levels—i.e. the estimate of the steady-state 10-year bond term premium, $\overline{r}^{\{40\}}-\overline{r}$, is added to the corresponding term-premium measured in deviation from its steady-state value, $-\frac{1}{40}\xi_t^{\{40\}}$, as described at the end of Section 2.2.

AL (thick line) model together with the ACM term premium (thin line) and the RE model (dashed line). Several conclusions emerge from this figure and the corresponding statistics shown in Table 4: AL term premium measures show smoother fluctuations and a milder downwards trend than the ACM term premium. These features of our estimated AL term premium are more in line with the findings in Bauer, Rudebusch and Wu (2014). They argue that term premia implied by affine term structure models are misleading because of small-sample bias, which alters the variability and trend behavior of risk-neutral rates. Thus, risk-neutral rates become significantly more volatile when accounting for small-sample bias in affine term structure models. Consequently, the implied term premia in affine models do not simply mirror the movements of long-term yields, but instead become more stable, which is consistent with our estimated AL term premia.

Moreover, Table 4 shows first- and second-moment statistics of the estimated term premia displayed in Figure 2. The mean of the baseline AL term premium is roughly half that of the ACM term premium (1.00 versus 2.09). Similarly, the fluctuations in the baseline AL term premium are milder than those of the ACM term premium (the standard deviations of these two term premium measures are 0.36 and 1.13, respectively). In spite of these different features, the two measures are highly contemporaneously correlated (0.85) and exhibit a high degree of persistence (the first-order autocorrelation coefficient is 0.96 for the ACM term premium and 0.88 for the baseline AL term premium).

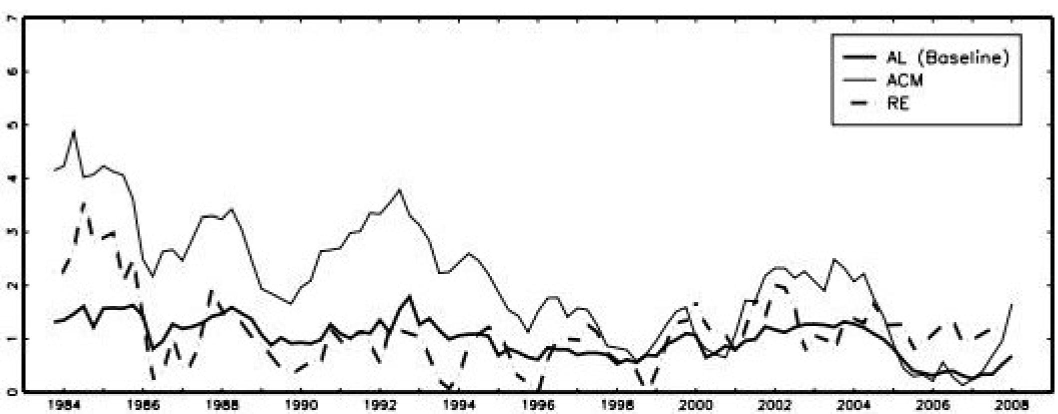


Figure 2. ACM, AL and RE (annualized) 10-year term premia

Note: The AL and RE term premia reported in this figure are computed as $\left[\left(\bar{r}^{\{40\}} - \bar{r}\right) - \frac{1}{40}\xi_t^{\{40\}}\right] \times 4$.

Comparing with RE, we find that the correlation of the alternative AL term premium measure with the ACM term premium measure is 0.85, whereas the estimated RE term premium shows a moderate correlation with the ACM and the AL term premium measures at 0.50 and 0.46, respectively. Hence, the contrast between the mild correlations under RE with the strong correlation between the ACM term premium and the AL term premium measure analyzed suggests that the AL model is able to do a much better job than RE in reproducing the fluctuations of the ACM term premium.

Additionally, the last row of Table 4 shows the correlation between the alternative term premium measures and the cyclical measure of GDP obtained from the Hodrick and Prescott filter (Hodrick and Prescott, 1997). The AL term premia and the ACM term premium are mildly countercyclical, in line with the findings in the related literature—among others, Campbell and Cochrane, 1999; Cochrane and Piazzesi, 2005; Bauer, Rudebusch and Wu, 2014. This finding stand in contrast with the acyclical RE term premium.

Table 4. Term premium statistics

	ACM	Baseline	RE
Mean	2.06	1.01	1.12
Std. deviation	1.11	0.36	0.67
First-order autocor.	0.96	0.88	0.80
Correlation matrix			
ACM	1.0		
Baseline	0.85	1.0	
RE	0.50	0.46	1.0
HP GDP	-0.36	-0.22	-0.01

5 Survey Data and expectations

PLM disciplined by the Survey of Professional Forecasters (SPF)

AL certainly introduces additional degrees of freedom, which may result in an arbitrary improvement in model fit—Adam and Marcet (2011). However, there is not much room for this type of criticism in our characterization of learning dynamics since we are considering a rather restrictive information set: only term structure information. Nevertheless, as a way of disciplining expectations, we assume that the deviations of agents' expectations on both inflation, $E_t \pi_{t+j}$, and consumption growth, $E_t \left(c_{t+j} - c_{t+j-1}\right)$, for j=1, 2, 3, 4 from the (observed) forecasts reported in the SPF follow AR(1) processes.Therefore, we add 8 new observables in the estimation exercises carried out in this section. In order to compare model fit with AL, a few versions of the RE model are also estimated disciplining RE with SPF forecasts.

Unlike Ormeño and Molnár (2015), we do not impose the more restrictive assumption that agents' expectations must match those of the SPF up to a white noise error. In short, we allow for persistent deviations between AL expectations and those reported in the SPF. This is because our extended model uses term structure information to characterize model expectations, which disciplines them in addition to SPF forecasts. As pointed out by Rudebusch and Williams (2009), Lahiri et al. (2013), and Stekler and Ye (2017), there is evidence that term structure information is not consistently used by professional forecasters—this is called "the yield spread puzzle" in the related literature. Thus, by allowing for the possibility of stationary but persistent deviations between AL expectations and those reported in the SPF we let market expectations through term structure data speak more freely when disciplining agents' expectations. Consequently, our estimated AL model expectations of inflation and consumption growth may transitorily deviate from their SPF counterparts, but broadly resemble their low-frequency features as shown later.

Figure 3 compares the four-quarter-ahead inflation and consumption growth in the model with those reported in the SPF. The graph at the top of this figure shows that the AL expectations generated by the model (medium-size line) reproduce the SPF (thick line) inflation downtrend. The model does a good job in reproducing the inflation forecasts reported in the SPF, but there are periods where the deviations between inflation expectations from the model and the SPF are rather persistent (e.g during the 1998-2004 period). As emphasized above, an important reason for these deviations of the inflation expectations in the model from SPF inflation forecasts is that the model's inflation expectations are driven by term structure information whereas SPF panelists do not seem to use this market information at all, as pointed out by Rudebusch and Williams (2009). This graph also shows inflation forecasts from the Michigan Consumer Survey. Clearly, SPF and model inflation forecasts are closer to each other than the Michigan Survey inflation forecasts. There are two reasons for this: First, Michigan inflation forecasts are not used to fit the model. Second, both SPF and the model consider GDP deflator inflation whereas the Michigan Survey considers CPI inflation.

The graph at the bottom of Figure 3 shows that the AL model does a great job in reproducing both the timing and the large low-frequency swings in consumption growth expectations reported in the SPF. This indicates that term structure information is closely in line with that used by SPF panelists when forecasting consumption growth.

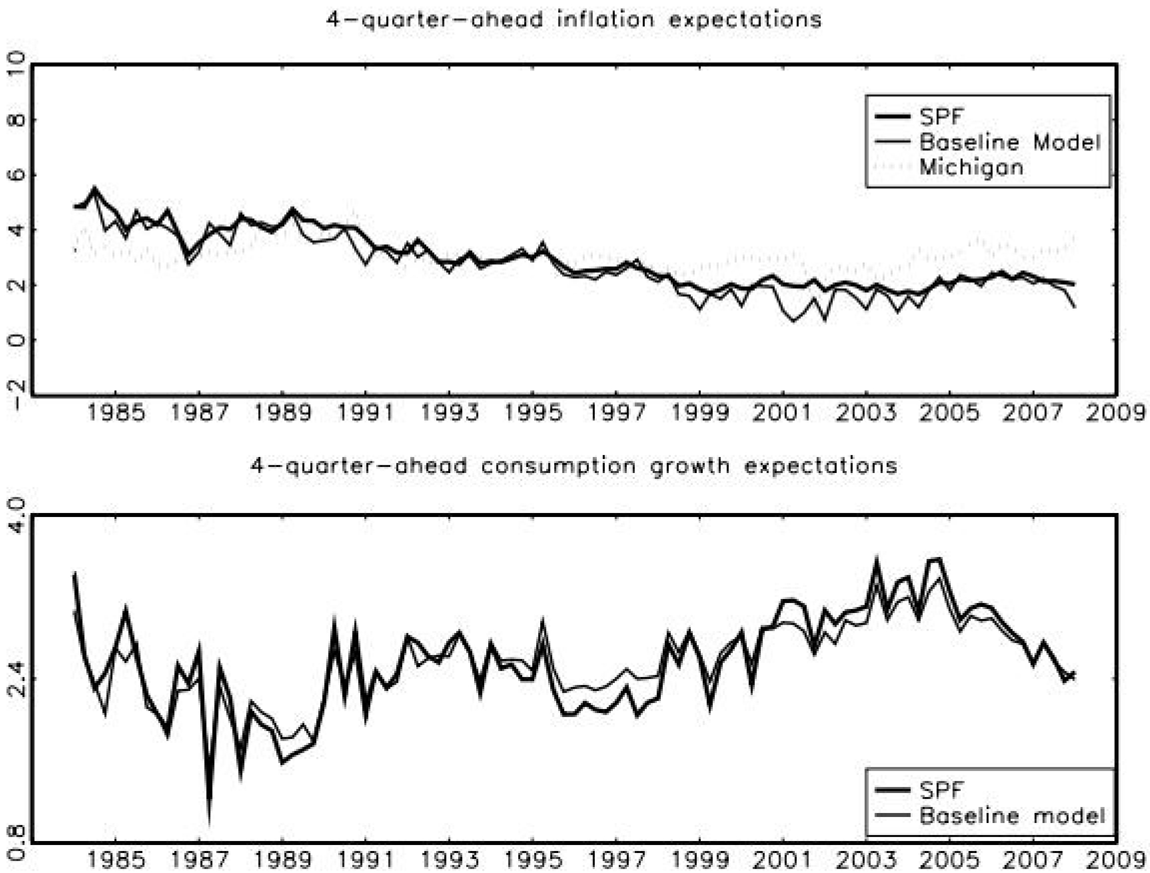


Figure 3. Model expectations versus SPF forecasts of inflation and consumption growth (annualized)

Finally, Table 5 shows the model fit in terms of the log data density across models when the SPF disciplines expectations. The longer the model horizon the larger the difference between AL and RE, which supports further the capacity of the term structure information in shaping expectations under AL.

Table 5. Model fit comparison: log data density

	ALTS-1Y	ALTS1-Y	RE1-Y	ALTS-10Y	ALTS-10Y	RE-10Y
	AR(1)	iid	AR(1)	AR(1)	iid	AR(1)
Observables	16	16	16	20	20	20
Log data density	218.73	-395.45	213.43	746.45	424.51	457.73

6 Conclusions

This paper shows how the term structure of interest rates contains important information to outperform traditional adaptive learning (AL) forecasting in DSGE when yields are priced using the standard consumption based asset pricing equation. Our AL model has the advantage of using simple forecasting rules that enables the term spread of interest rates to fully characterize the expectations of all forward-looking variables.

The estimation results show that the extended AL model with term structure provides a much better fit to the data than the rational expectations version. This finding is made very clear when the whole yield curve is fitted. The rationale behind this result is the additional persistence introduced from adding multi-period Euler equations associated with the pricing of different bond yields under AL. This approach enhances the expectation channel in the model, standing in sharp contrast to RE, where multi-period forecasting is redundant for agents expectations. In addition, our AL model generates a term premium, which is highly correlated with term premium measures obtained from no-arbitrage affine term structure models (e.g. Adrian, Crump and Moench, 2013). Finally, we show that our results are robust to the incorporation of survey data to disciplining AL expectations. Our model implied expectations are also succesful in replicating the business cycle behavior of expectations reported in survey data.

References

- Adam, Klaus, and Albert Marcet. 2011. "Internal rationality, imperfect market knowledge and asset prices." *Journal of Economic Theory* 146, 1224-1252.

- Adrian, Tobias, Richard K. Crump, and Emanuel Moench. 2013. "Pricing the term structure with linear regressions." *Journal of Financial Economics* 110, 110-138.

- Aguilar, Pablo A., and Jesús Vázquez. 2021. "An estimated DSGE model with learning based on term structure information." *Macroeconomic Dynamics, 25(7).*

- Bauer, Michael D., Glenn D. Rudebusch, and Jing C. Wu. 2014. "Term premia and inflation uncertainty: Empirical evidence from an international panel dataset: Comment." *American Economic Review* 104, 323-333.

- Bhansali, Rajendra J. 2002. "Multi-step forecasting," in Michael P. Clements and David F. Hendry, eds., *A Companion to Economic Forecasting*. Oxford: Blackwell Publishing, pp. 206-221.

- Campbell, John Y., and John H. Cochrane. 1999. "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy* 107, 205-251.

- Cochrane, John H., and Monika Piazzesi. 2005. "Bond risk premia." *American Economic Review* 95, 138-160.

- Croushore, Dean. 2011. "Frontiers of real-time data analysis". *Journal of Economic Literature* 49, 72-100.

- Crump, Richard K., Stefano Eusepi, and Emanuel Moench. 2017. "The term structure of expectations and bond yields." Federal Reserve Bank of New York Staff Reports, no.775.

- De Graeve, Ferre, Marina Emiris and Rafael Wouters. 2009. "A structural decomposition of the US yield curve." *Journal of Monetary Economics* 56, 545-559.

- Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti. 2017. "Safety, liquidity, and the natural rate of interest." Federal Reserve Bank of New York, *mimeo.*

- Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework." *Econometrica* 57, 937-969.

- Estrella, Arturo, and Gikas A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46, 555-576.

- Estrella, Arturo, and Frederic S. Mishkin. 1997. "The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank." *European Economic Review* 41,1375–1401.

- Evans, George W., and Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.

- Eusepi, Stefano, and Bruce Preston. 2011. "Expectations, learning, and business cycle fluctuations". *American Economic Review* 101, 2844-2872.

- Eusepi, Stefano, and Bruce Preston. 2018. "Fiscal foundations of inflation: Imperfect knowledge". *American Economic Review* 108, 2551-2589.

- Greenwood, Robin, Samuel G. Hanson, and Jeremy C Stein. 2015. "A comparative-advantage approach to government debt maturity." *Journal of Finance* 70, 1683-1722.

- Gürkaynak, Refet. S., Brian Sack, and Jonathan H. Wright. 2007. "The U.S. Treasury yield curve: 1961 to the present." *Journal of Monetary Economics* 54, 2291-2304.

- Hodrick, Robert J., and Edward C. Prescott. 1997. "Postwar U.S. business cycles: An empirical investigation." *Journal of Money, Credit, and Banking* 29, 1-16.

- Hommes, Cars, and Gerhard Sorger. 1998. Consistent expectations equilibria. *Macroeconomic Dynamics* 2, 287-321.

- Honkapohja, Seppo, Kaushik Mitra, and George W. Evans. 2012. "Notes on agents' behavioral rules under adaptive learning and studies of monetary policy". CDMA11/02 working paper, University of St Andrews.

- Jordà, Òscar. 2005. "Estimation and inference of impulse responses by local projections." *American Economic Review* 95, 161-182.

- Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. "The aggregate demand for Treasury debt." *Journal of Political Economy* 120, 233-267.

- Lahiri, Kajal, George Monkroussos, and Yongchen Zhao. 2013. "The yield spread puzzle and the information content of SPF forecasts." *Economics Letters* 118, 219-221.

- Marcet, Albert, and Thomas J. Sargent. 1989. "Convergence of least-squares learning in environments with hidden state variables and private information." *Journal of Political Economy* 97, 1306–1322.

- McCallum, Bennett T. 1994. "Monetary policy and the term structure of interest rates". NBER Working Paper 4938. Reprinted in the *Federal Reserve Bank of Richmond Economic Quarterly*, Volume 91/4 Fall 2005.

- Milani, Fabio. 2007. "Expectations, learning and macroeconomic persistence." *Journal of Monetary Economics* 54, 2065-2082.

- Milani, Fabio. 2008. "Learning, monetary policy rules, and macroeconomic stability." *Journal of Economic Dynamics and Control* 32, 3148-3165.

- Milani, Fabio. 2011. "Expectation shocks and learning as drivers of the business cycle." *Economic Journal* 121, 379-401.

- Mishkin, Frederic S. 1990. "The information in the longer maturity term structure about future inflation." *Quarterly Journal of Economics* 105, 815-828.

- Ormeño, Arturo, and Krisztina Molnár. 2015. "Using survey data of inflation expectations in the estimation of learning and rational expectations models." *Journal of Money, Credit, and Banking* 47, 673-699.

- Orphanides, Athanasios, and John C. Williams. 2005. "The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations." *Journal of Economic Dynamics and Control* 29, 1927-1950.

- Preston, Bruce. 2005. "Learning about monetary policy rules when long-horizon expectations matter." *International Journal of Central Banking* 1, 81-126.

- Rudebusch, Glenn D., and Eric T. Swanson. 2012. "The bond premium in a DSGE model with long-run real and nominal risks". *American Economic Journal: Macroeconomics* 4, 105–143.
- Rudebusch, Glenn D, and John C. Williams. 2009. "Forecasting recessions. The puzzle of the enduring power of the yield curve." *Journal of Business Economics and Statistics* 27, 492-503.
- Sargent, Thomas. J. 1991. Equilibrium with signal extraction from endogenous variables. *Journal of Economic Dynamics and Control* 15, 245-273.
- Sinha, Arunima. 2015. "Government debt, learning and the term structure." *Journal of Economic Dynamics and Control* 53, 268-289.
- Sinha, Arunima. 2016. "Learning and the yield curve." *Journal of Money, Credit, and Banking* 48, 513-547.
- Smets, Frank R., and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review* 97, 586-606.
- Slobodyan, Sergey, and Rafael Wouters. 2008. "Estimating a medium-scale DSGE model with expectations based on small forecasting models." *Mimeo.*
- Slobodyan, Sergey, and Rafael Wouters. 2012a. "Learning in a medium-scale DSGE model with expectations based on small forecasting models." *American Economic Journal: Macroeconomics* 4, 65-101.
- Slobodyan, Sergey, and Rafael Wouters. 2012b. "Learning in an estimated medium-scale DSGE model." *Journal of Economic Dynamics and Control* 36, 22-46.
- Slobodyan, Sergey, and Rafael Wouters. 2017. "Adaptive learning and survey expectations of inflation." National Bank of Belgium *mimeo*.
- Stekler, Herman O., and Tianyu Ye. 2017. "Evaluating a leading indicator: an application—the term spread". *Empirical Economics* 53, 183-194.
- Vázquez, Jesús. 2019. "The bond term premium in an estimated DSGE model with learning." University of the Basque Country (UPV/EHU) *mimeo*.

- Vázquez, Jesús and Aguilar Pablo A. 2021. "Adaptive learning with term structure information," *European Economic Review,* vol. 134.

- Vázquez, Jesús, Ramón María-Dolores, Juan M. Londoño. 2013. "On the informational role of term structure in the US monetary policy rule." *Journal of Economic Dynamics and Control* 37, 1852-1871.

Part III

Can news help measure economic sentiment? An application in COVID-19 times

1 Introduction

Benchmark data to assess activity are normally provided by the national accounts and are available at a quarterly frequency. Nevertheless, there usually exists a significant publication lag of, typically, at least 30 days after the quarter of reference has ended. More timely data is usually published in the form of economic indicators, both covering quantitative information (so-called "hard" indicators) and qualitative information provided by agent surveys of sentiment and plans (so-called "soft" indicators). These standard leading indicators, used by practitioners and academics alike, are typically available with a short delay, and are available at a monthly frequency (see Giannone et al., 2009 or Keeney et al., 2012). Technological progress has enabled the development of other sources of data usable for monitoring real-time economic activity, in particular, indicators based on textual analysis, including news media (see Shapiro et al., 2020, Thorsrud, 2016, 2020 and many others). Following this line of research, this paper proposes a new daily economic news sentiment indicator (DENSI) based on newspaper data for Spain as an alternative measure of economic confidence, which can be constructed at daily frequencies and it is shown to better help the nowcasting of Spanish GDP growth than a widely used popular confidence indicator —i.e. the Economic Sentiment

With Corinna Ghirelli, Matías Pacce and Alberto Urtasun (BdE), published in Economics Letters (2021), Vol. 199

Indicator (ESI)[1] published by the European Commission (EC)—. To construct our measure we rely on Boolean searches,[2] which boils down to counting the frequencies of specific words in the text, a quite popular procedure in the recent literature: for instance, see the influential economic policy uncertainty (EPU) index by Baker et al. (2016).[3]

Our indicator, as it is the case for many other textual based indicators, could potentially be interpreted as a measure of economic confidence and, therefore, compared with other confidence indicators, which are mostly survey-based. Survey-based confidence indicators are probably among the most useful soft indicators available to make an assessment of the general economic situation. They reflect agents' perceptions regarding the present and future economic situation. Indeed, for each particular country, confidence indicators are highly correlated with GDP growth (e.g. see the ESI). However, the COVID-19 crisis has put some of these indicators' weaknesses in the spotlight.

First, most confidence indicators rely on surveys that are conducted during (at most) the first part of the reference month (for instance the ESIs surveys are carried out in the first three weeks of the reference month),[4] while results are released during the last week of that month. In general, this strategy would not cause any problem, but if a significant shock occurs during the second half of the month, the confidence indicators would not reflect it (Shapiro et al., 2020).

Second, as for any survey, its accuracy largely depends on the survey's responses rates and, when these are not high enough, potential sampling problems can emerge, especially during some particular situations (Ludvigson, 2004). Both of these problems have simultaneously emerged during the COVID-19 outbreak, given it started to affect most European countries

[1]The ESI is built as a composite of five different confidence indicators using the following weights: 40% for industry, 5% for construction, 30% for services, 20% for consumers, 5% for retail

[2]Alternatively, it would be possible to rely on natural language processes in order to extract the "sentiment" or the topics characterizing the news (e.g., Shapiro et al., 2020, Thorsrud, 2020, 2016).

[3]Following this procedure, a number of EPU indexes have been developed for many countries (e.g. Ghirelli et al., 2019 for Spain) and has been used in a number of empirical applications (e.g. Colombo, 2013, Caggiano et al., 2017, Fontaine et al., 2017, Meinen and Roehe, 2017).

[4]For example, during the month of March 2020, the ESI surveys of consumers, the service sector, and the retail sector were carried out between the 2nd and the 12th of March, while the surveys of the industry sector and the construction sector were carried out between the 2nd and the 23rd of March

during the second half of March and due to the difficulty of carrying out the surveys during the lockdown. This situation became evident in the ESI release for March 2020. For instance, the value of the ESI for March did not reflect the pandemic crisis outbreak that started in the middle of the month,[5] but instead remained at the optimistic levels reached in the previous months (see the March 2020 observation in the solid line in Figure 2). Moreover, for the case of Italy, the EC announced that no data could be collected for the month of April due to the strict confinement measures.

A third issue is that survey-based confidence indicators can show some specific bias along the business cycle. According to Gayer and Marc (2018), this could happen, for example, if business managers and consumers adapt their economic expectations to a more modest growth. Such a bias became evident at the end of the Great Recession, when average economic growth was significantly lower than that observed during the pre-crisis period but the ESI stood at similar average levels in both periods.[6] This bias is evident in Figure 2: from 2015, the ESI frequently had values in the 75–90th percentile of its own distribution, probably being overoptimistic. Figure 2 also confirms that this bias is not present in our proposed indicator. Its dynamics, instead, correctly reflect the moderate growth of the GDP during the last recovery, as opposed to that observed in the pre-crisis period. A reason for the presence of this bias in survey-based indicators may be the possibility that respondents base their perceptions of the economic situation on a cross-section comparison, i.e., ignoring the historical perspective. In contrast, when journalists write an article about the economic situation of a country, they typically make an historical assessment, which may explain why our news-based indicator does not show such a bias.

Dealing with these weaknesses, our proposed indicator based on newspaper articles emerges as an alternative confidence indicator, especially since the COVID-19 pandemic crisis. Newspaper articles may be a useful data source for constructing confidence indicators to assess

[5]Spain was locked down on the 14th of March 2020.

[6]The average annual GDP growth rate between 2000 and 2007 was around 3.7% and diminished to 2.6% during the 2014–2019 period. Analyzing the same periods, the average levels for the ESI were, respectively, 105.5 and 106.3 points.

the current economic situation, for two main reasons: first, the press (especially financial newspapers or the economic sections of generalist newspapers) provides a daily picture of the economic developments of a country; second, since newspapers are published every day, this allows constructing real-time indicators at high (daily) frequencies. Finally, text-based sentiment measures can be developed at a very low cost relative to survey-based measures.

When proposing the new confidence indicator, we will focus on the Spanish case. Therefore, we rely on a large database of Spanish press and construct a sentiment indicator that reflects the economic tone of news articles in the short-run. In a nutshell, our indicator reflects the balance between the number of news articles that contain keywords related to upturns and downturns of the Spanish business cycle. The new proposed indicator closely follows the dynamics shown by the ESI. Nonetheless, in contrast to the ESI, in March 2020 the DENSI drops to the level of the Great Recession, which appears to be a more correct description of the economic outlook given the start of the pandemic crisis.

In order to analyze whether the DENSI could do a better job than survey-based confidence indicators regarding the assessment of the current economic situation, we perform two different—but related—exercises. First, relying on mixed-frequency bivariate vector autoregressive (MF-BiVAR) models, we show that a model including both the DENSI and GDP significantly improves the GDP forecast accuracy when compared to a model that includes the ESI and GDP as the main variables. This result suggests that the DENSI could provide a better signal for an earlier evaluation of the economic situation than the ESI. Second, we use both indicators to compute the implied probability of recessions at short-term horizons. Result show that the DENSI does better than the ESI in predicting the business cycle turning points in the short-run.

A number of papers have constructed text-based measures of economic activity. The latter are becoming increasingly popular because they are real time measures (Shapiro et al., 2020). E.g., Calomiris and Mamaysky (2019) use news articles to predict risk and return in stock markets, while Fraiberger (2016) constructs a sentiment index based on the economic news articles produced by Reuters. Nyman et al. (2018) exploit financial market text-based data to predict financial system distress. Thorsrud (2020, 2016) show that the information contained in the main Norwegian business newspapers articles improves the GDP nowcast. Our paper is closelyrelated to Combes et al. (2018) and Shapiro et al. (2020). The first authors, who focused on France, construct a sentiment indicator of economic activity based on news published in Le Monde, the main generalist French newspaper, and show that the index helps to improve the short-term GDP forecast and nowcast. In particular, their textual indicator improves forecasting compared to a simple autoregressive model and an autoregressive model augmented with the survey-based Insee Business Climate indicator.[7] Shapiro et al. (2020) develop an indicator of economic sentiment based on economic and financial newspaper articles and show that daily news sentiment is predictive of movements of survey-based measures of consumer sentiment. In particular, they estimate impulse responses of the main macroeconomic variables to shocks in economic sentiment and find that positive shocks in economic sentiment increase consumption, output, and interest rates, whereas they dampen inflation. Therefore, they claim that newspaper-based sentiment indicators provide useful signals of economic activity at a very low cost relative to survey-based measures.

We contribute to this literature by proposing a new newspaper-based sentiment indicator for Spain, which could be obtained on a daily basis at any point during the month. Moreover, we provide evidence that newspaper-based indicators can better deal with the aforementioned limitations, which are typical of survey-based confidence indicators (i.e., (i) the monthly value refers only to the first half of each month; (ii) potential sampling problems due to a lower response rate; (iii) potential bias along business cycles). Finally, in line with the previously mentioned literature, we also show that newspaper-based confidence indicators can improve classical confidence indicator measures when assessing more recent economic developments. The rest of the paper is organized as follows. In Section 2, we describe our DENSI. Section 3 presents two empirical applications in which we compare the predictive power of the ESI

[7] This indicator is equivalent to the ESI for France.

and the DENSI. Section 4 offers some concluding remarks.

2 Description of the index

We build a sentiment indicator capturing the economic tone of news articles published in the Spanish press. In a nutshell, our indicator reflects the balance between the number of news articles that contain keywords related to upturns and downturns in the Spanish business cycle.

We consider 7 relevant Spanish national newspapers: El País, El Mundo, La Vanguardia, ABC, Expansión, Cinco Días, and El Economista. The first 4 newspapers are the largest and most-read generalist newspapers in Spain, while Expansión, Cinco Días, and El Economista are the three main Spanish business newspapers. We focus on the printed editions of these newspapers and ignore the online versions. Our decision is supported by two arguments: first, online editions have become more widespread only in the most recent years, while we consider quite a substantial time span. Ignoring the digital press ensures homogeneity in our index. Second, printed newspapers ensure meeting standards in terms of the quality and relevance of the news since editors select the articles to be published in the printed version of the newspaper given space limitations.

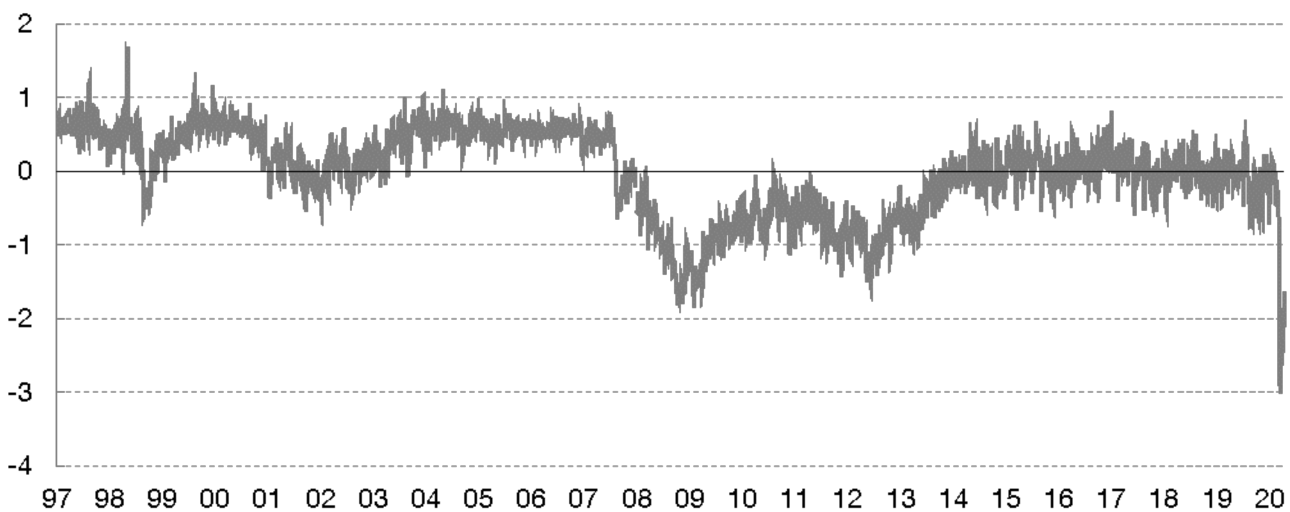
Our main indicator is monthly. However, one advantage of relying on newspapers articles is that one can construct indicators at higher (daily) frequencies. All searches are carried out using the Dow Jones' Factiva service. For each newspaper, we conduct queries from the first month the newspaper is collected in the Dow Jones' Factiva database, starting from January 1997. We restrict all queries to the following articles: (i) articles in the Spanish language; (ii) articles 6with content related to Spain, based on Factiva's indexation; (iii) articles about corporate or industrial news, economic news, or news about commodities or financial markets, according to Factiva's indexation. We then perform three types of queries. First, we count the number of articles that satisfy the aforementioned requirements, which provides us with a measure of economic articles. This will serve as denominator for our indicator. Second, we count the number of articles that, in addition to satisfying the aforementioned conditions, contain upswing-related keywords. That is, the articles must contain the word recuperación* (recovery) or one of the following words, provided that they are preceded or followed by either economic* (economic) or economía (economy) within a distance of 5 words: aceler* (acceleration), crec*, expansi* (growth), increment*, aument* (increase), recuper* (recovery), mejora* (improvement). In addition, in order to ensure that the news items are about the Spanish business cycle, we also require the article to contain the word Españ* (Spain/Spanish).[8] Similarly, we count the number of articles that, in addition to satisfying the aforementioned conditions, are about downswings. In particular, the articles must contain the word recesión* (recession) or crisis (crisis) or one of the following words, provided that they are preceded or followed by either economic* (economic) or economía (economy) within a distance of 5 words: descen*, disminu*, redu* (decrease), ralentiz*, decrec*, desaceler*, contracci´on* (slowdown). The articles should also contain the word Españ* (Spain/Spanish). Then, for each newspaper, we take the difference between the upturn and downturn counts.

To construct the index, we follow the procedure used by Baker et al. (2016). First, we scale the difference between upturn and downturn counts by the total number of economic articles in the same newspaper/month. Second, we standardize the monthly series of scaled counts considering the period of 1997m1-2020m2 so that the volatility of each series is comparable across newspapers. Third, we average the series across newspapers. Fourth, we rescale the resulting index to mean 0, considering the period of 1997m1-2020m2, to obtain the monthly DENSI.

[8] We already require that the content of the article is related to Spain. However, the latter condition is quite loose since articles that are related to Spain do not necessarily talk about Spanish matters (e.g. they may talk about foreign matters that have implications for Spain). For this reason, we also require that the article contains the word Spain/Spanish.

Figure 1 shows the 7-day moving average of the daily DENSI. A major advantage of newspaper based indexes compared to the traditional monthly confidence indicators is that each day they incorporate new information in the construction of the index. This becomes crucial when important shocks occur in the middle of the month or far from the date at which the monthly confidence indicators are released, i.e. when Spain was locked down on the 14th of March due to the COVID-19 outbreak.

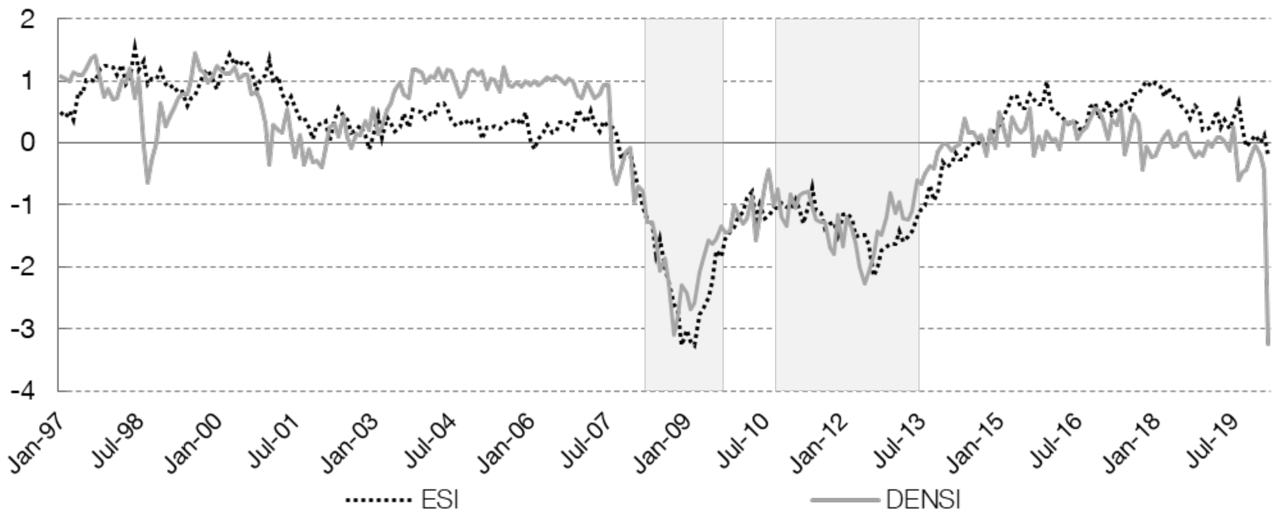
Figure 1: The DENSI at a daily frequency



Note: The line depicts the 7-day moving average of the DENSI index.

Figure 2 shows our monthly DENSI against the ESI. Both indexes are highly correlated. However, their behavior differs dramatically in March 2020 after the COVID-19 outbreak in Spain. The DENSI signals the outbreak correctly and drops to values reached during the Great Recession, while the ESI remains at the positive levels characterizing the recent months. This is because the ESI relies on answers mostly collected in the first half of the month, which do not capture the start of the economic crisis that occurred from the lockdown implemented on March 14th to fight against the spread of COVID-19. In the following section we provide a case-study of the COVID-19 recession anticipated using this indicator

Figure 2: Comparison between the ESI and the DENSI



Note: The gray area represents recessions according to the Spanish Business Cycle Dating Committee. The ESI is normalized to have the same range as the DENSI

3 Empirical analysis

The ESI has been shown to be helpful in predicting short-term GDP growth (e.g., see Cervena and Schneider, 2014, Mazzi et al., 2014, Gajewski, 2014). In this section, we aim to go one step further by evaluating whether our proposed indicator could do better than the ESI in this sense Figure 2: Our indicator at a daily frequency Note: The line depicts the 7-day moving average of the DENSI index. and, consequently, we present two different but related exercises to analyze if this is the case. In the first one, we rely on the ESI and the DENSI to forecast Spanish GDP growth and we compare the forecast accuracy obtained in each case. In the second application, we set up a model to compute the implied probability of recession for Spain based on, alternatively, the ESI and the DENSI. As in the previous exercise, the main objective is to compare the predictive power of each indicator.

3.1 Forecasting GDP

GDP figures, which are probably the best measure to proxy the aggregate state of the economy, are generally released at quarterly frequencies. Of course, between each quarterly publication, many meaningful economic indicators become available that help in the assessment of the economic situation. Therefore, mixed-frequency models emerge as one way of exploiting this intra-quarter information when forecasting short-term GDP growth. Here, we rely on an MFBiVAR model, as in Mariano and Murasawa (2010), to compare the GDP forecast accuracy of a model that alternatively includes the ESI or the DENSI as monthly variables (obviously, GDP is the quarterly variable included in both cases).

In Mariano and Murasawa's (2010) notation, the GDP quarter-on-quarter growth rate (y_t), which is observed every three periods (months), could be written in terms of a latent monthly growth rate (y_t^*) through the following arithmetic formula (see Mariano and Murasawa, 2010 for its derivation):

$$y_t \;=\; \frac{1}{3}y_t^* + \frac{2}{3}y_{t-1}^* + y_{t-2}^* + \frac{2}{3}y_{t-3}^* + \frac{1}{3}y_{t-4}^* \; . \tag{1}$$

Of course, y_t^* is never observed. By letting x_t stand for the monthly variable (the ESI or the DENSI), it is possible to define $Y_t = (y_t, x_t)'$ and $Y_t^* = (y_t^*, x_t)'$ and relate both vectors as[9]

$$Y_t \;=\; H(L)Y_t^* \; , \tag{2}$$

where

$$H(L) \;=\; \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2/3 & 0 \\ 0 & 0 \end{pmatrix} L + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} L^2 + \begin{pmatrix} 2/3 & 0 \\ 0 & 0 \end{pmatrix} L^3 + \begin{pmatrix} 1/3 & 0 \\ 0 & 0 \end{pmatrix} L^4 \; ,$$

where L is a lag operator. As indicated in Mariano and Murasawa's (2010), a Gaussian vector

[9] All variables are standardized to have a zero mean and a variance equal to one before estimating the model.

autoregressive of order P (VAR(P)) model could be assumed for Y_t^* as

$$\Phi(L)Y_t^* = w_t, \tag{3}$$

where $w_t \sim IN(0, \Sigma)$. For $P \leq 5$ and defining $S_t = \left(Y_t^*, ..., Y_{t-4}^*\right)'$, the mixed-frequency VAR(P) model could then be written under the state-space representation as

$$\begin{aligned} Y_t &= CS_t \ , && (4) \\ S_{t+1} &= AS_t + Bz_t \ , && (5) \end{aligned}$$

where $z_t \sim IN(0, I_2)$ and

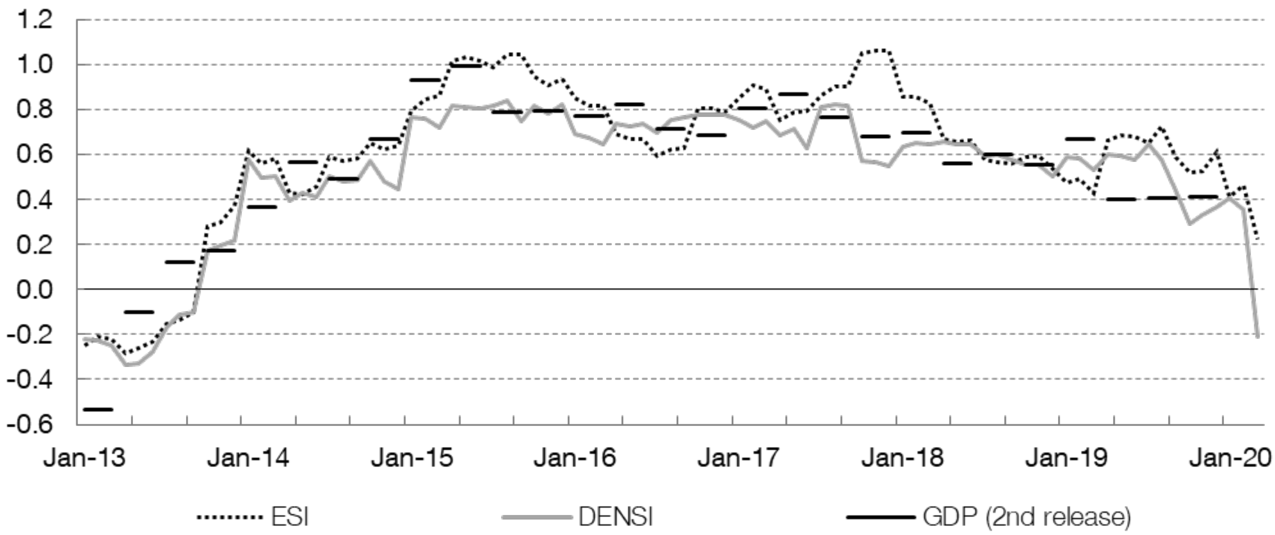
$$A = \begin{pmatrix} \phi_1 & \cdots & \phi_p & 0_{2\times(5-p)2} \\ & I_8 & & 0_{8\times 2} \end{pmatrix}, \qquad B = \begin{pmatrix} \Sigma^{1/2} \\ 0_{8\times 2} \end{pmatrix}, \qquad \text{and} \qquad C = \begin{pmatrix} H_0 & \cdots & H_4 \end{pmatrix}.$$

The previously described state-space model is estimated by means of maximum likelihood estimation. Under this specification (we assume $P = 3$), we carry out the following forecasting exercise.

First, we set up a pseudo-real-time nowcasting exercise for GDP. In particular, we assume it is the last day of each month t and, for our exercise, we use the information that would have been available at that moment (we use real-time vintages for GDP). It should be noted that flash estimates for GDP are published 30 days after the end of the reference quarter. Therefore, within a quarter, we predict the one-quarter-ahead GDP growth no matter whether it is the first, second, or third month of the quarter. We start the exercise in January 2013, and for each month until December 2019, we conduct the nowcast of quarter-on-quarter GDP growth when the ESI and the DENSI are, alternatively, the monthly variable included in the model. Our target is the second release of GDP figures. The nowcasts are shown in Figure 3. The dashed and solid lines represent the predictions based on the model with the ESI and

DENSI, respectively, while the horizontal line is the target value (second release of GDP).

Figure 3: Quarterly GDP growth nowcast



Note: The dashed (solid) line represents the predictions obtained from the model that includes the ESI (DENSI) and the GDP. The short line depicts the target variable (second release of GDP).

To compare the forecast accuracy of both models, we rely on the forecast root mean squared error (RMSE). Results indicate that the RMSE of the Bi-VAR model that contains the DENSI is about 20% lower than that obtained under the model that includes the ESI, the difference being significant at the 1% confidence level.11 This indicates that the DENSI significantly improves the nowcast of GDP when compared with the ESI, meaning that it provides a potentially better signal for the evaluation of the current economic situation, being a clear example the prediction for the beginning of 2020.

3.2 Predicting the probability of recession

In our second application, we test whether our text-based sentiment indicator has predictive content in terms of business-cycle turning points. We start from a very standard model, where recession probabilities at future horizons are estimated as a (probit) function of the present value of the slope of the yield curve as in. Wright (2006). To measure the stance of the business cycle, we rely on the recession dates given by the Spanish Business Cycle Dating Committee (Spanish Economic Association, 2015). Departing from this model, we check whether forecast accuracy improves by adding alternatively the DENSI or the ESI as an additional regressor. We estimate these models up to a 12-month horizon considering the period from January 1997 until February 2020. Then we predict recession probabilities using both models during the COVID-19 outbreak. In particular, we estimate the following models

$$P(R)_{(t+n)} = \alpha + \beta_1 \left[yield_t^{3M} - yield_t^{10A}\right] + \epsilon_{(t+n)}, \tag{Model 1}$$

where $P(R)_{(t+n)}$ stands for the recession probability at a future horizon $(t+n)$, $yield_t^{3M}$ and $yield_t^{10A}$ are, respectively, the three-month and 10-year Spanish treasury bill yields, and the residuals ϵ_t are assumed to be independently and normally distributed. To measure the stance of the business cycle, we rely on the recession dates given by the Spanish Business Cycle Dating Committee (Spanish Economic Association, 2015), and we assume that the recession probability is equal to one when a recession occurred and zero otherwise. Departing from this model, we check whether forecast accuracy could be improved by adding an economic sentiment indicator (the DENSI or the ESI) as an additional regressor. Therefore, we estimate the following models, which encompass Model 1:

$$P(R)_{(t+n)}^{ESI} = \alpha^{ESI} + \beta_1^{ESI} \left[yield_t^{3M} - yield_t^{10A}\right] + \beta_2^{ESI} ESI_t + \epsilon_{(t+n)}^{ESI}, \tag{Model 2}$$

$$P(R)_{(t+n)}^{DENSI} = \alpha^{DENSI} + \beta_1^{DENSI} \left[yield_t^{3M} - yield_t^{10A}\right] + \beta_2^{DENSI} DENSI_t + \epsilon_{(t+n)}^{DENSI}. \tag{Model 3}$$

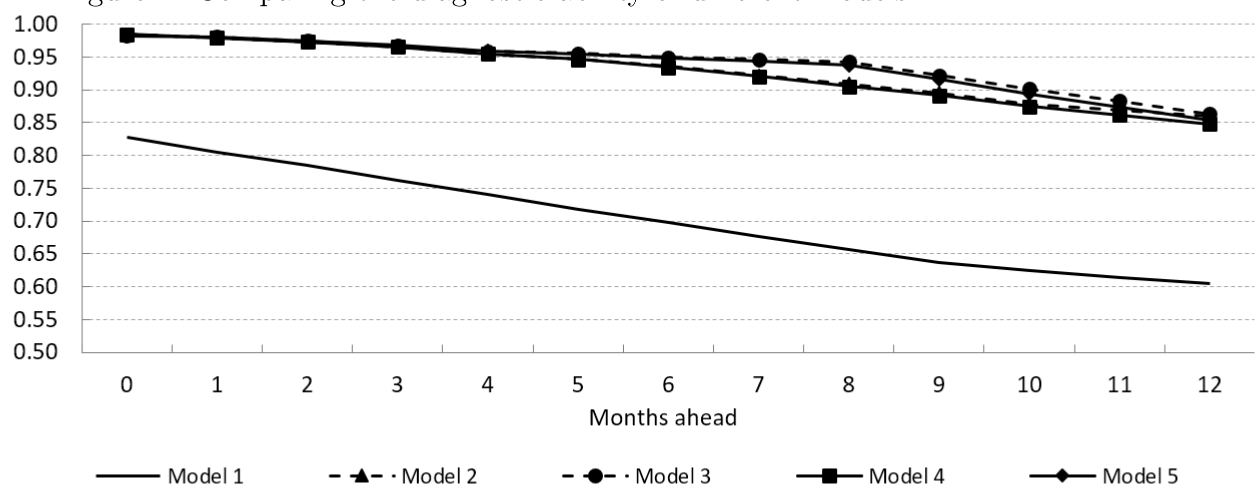
Finally, we compute the recession probabilities when the slope of the yield curve is omitted from the model:

$$P(R)^{ESI^*}_{(t+n)} = \alpha^{ESI^*} + \beta_2^{ESI^*} ESI_t + \epsilon^{ESI^*}_{(t+n)}, \qquad \text{(Model 4)}$$

$$P(R)^{DENSI^*}_{(t+n)} = \alpha^{DENSI^*} + \beta_2^{DENSI^*} DENSI_t + \epsilon^{DENSI^*}_{(t+n)}. \qquad \text{(Model 5)}$$

In figure 4 shows the ROC curve under each particular model. Three points are worth noting. First, in the case of Spain, the standard model (Model 1) is outperformed by all of the other four selected models at all horizons (the blue line is always below the other lines). The low performance of the slope of the yield curve as a predictor of recession could potentially be explained by the key role played by the European Central Bank since the 2012 financial crisis.[10] In particular, conventional and unconventional monetary policies such as forward guidance, asset purchase, and enhanced credit support may have flattened the slope of the yield curve associated with national bonds, curbing the predictive power of the indicator. Second, the models that include either the ESI or the DENSI show a great fit in terms of ROC, being over 0.9 at any horizon, with the DENSI having slightly better diagnostic ability at longer horizons. Third, once any of the confidence indicators is included in the model, adding the yield curve as a regressor does not improve the results.

Figure 4: Comparing the diagnostic ability of different models

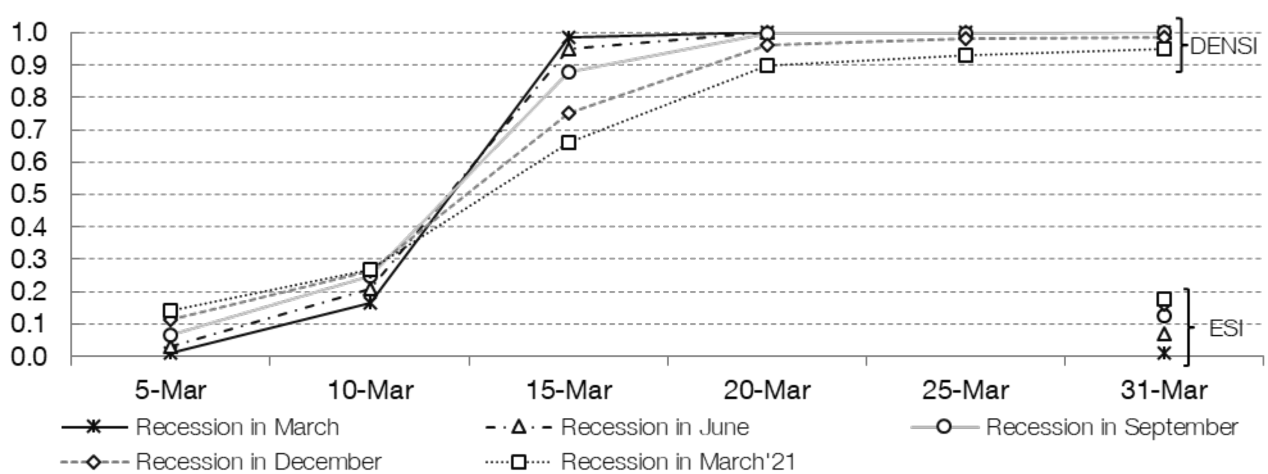


[10]Based on the period before the Great Recession, Alonso et al. (1997) and Mart ́ınez Serna and Navarro (2005) show that the inclusion of the Spanish yield curve provides relevant information to predict Spanish output slowdowns or expected economic growth, respectively

Note: Each line represents the receiver operating characteristic (ROC) curve of each model. A value of 1 represents a perfect ability to discriminate between two alternative statuses (recession or not), whereas a value of 0.5 represents no ability of discrimination. Model 1: slope of the yield curve; Model 2: slope of the yield curve and ESI; Model 3: slope of the yield curve and DENSI; Model 4: ESI; Model 5: DENSI.

Finally, Figure 5 shows the estimated probabilities of recession from March 2020 up to 3, 6, 9, and 12 months ahead using only either DENSI or ESI (models 4 and 5). To show the importance of having real-time updates of the DENSI, we evaluate the model using the information available on the 5th, 10th, 15th, 20th, 25th, and 31st of March. On each date, we assume that the value of the DENSI for March is the average value of the already known days of the month. Therefore, the value computed on the 31st of March matches the monthly value of the DENSI, while the values computed on earlier dates are approximations of that monthly value. As for the ESI, there is only one date available at the end of the month.

Figure 5: Recession probabilities during the COVID-19 outbreak



Note: The x-axis represents the dates in March 2020 when we compute the probability of recession at different horizons based on the model with the DENSI: 3/2020 (blue squares), 6/2020 (purple squares), 9/2020 (yellow squares), 12/2020 (red squares), 3/2021 (brown squares). Recession probabilities obtained with the model with the ESI are represented by the circles only at the end of the month.

The estimations based on the DENSI adjust correctly to the rapid economic developments due to the COVID-19 outbreak. By March 14th, when the country was immediately locked down,[11] the probability of recession increased, at least, by a factor of two at any horizon, and reached values near 1 during the second half of the month. Since most of the ESI's surveys took place before the lockdown, it is not surprising that the recession probability for March based on the ESI remains very low. These results indicate that in the presence of big events, the DENSI would be able to capture their impact earlier than the ESI.

4 Conclusions

For any policymaker, a prompt assessment of the current economic situation is of key importance. Since most macroeconomic variables are feasible only after some lag, having meaningful up-to date indicators of the state of the economy becomes a priority. Therefore, we have proposed a new indicator of economic sentiment based on newspaper articles that allows assessing real-time economic activity in Spain. Our proposed indicator has three major advantages with respect to the very popular ESI, which is provided by the European Commission: 1) it can be constructed on a daily basis; 2) it does not suffer from bias along business cycles; and 3) it is very flexible with respect to the topic of interest (one can easily develop new indicators focused on specific issues by selecting specific keywords of interest and constructing the indicators backwards, whereas with survey-based data, new topics need the development of new questions in the survey).

In the empirical part of the paper, we set up two exercises to show that the DENSI is a better indicator than the ESI for the case of Spain. In a first exercise, we set up mixed frequency bivariate VAR models and let the ESI and the DENSI compete in nowcasting the GDP growth rate. The DENSI provides significantly more precise predictions from January 2013 to December 2019. In addition, we showed two examples in which a major shock occurred 18in the middle of the month (in September 2008, with the Lehman Brothers

[11] According to Article 116.2 of the Spanish Constitution, a state of alarm can be declared "in all or part of the national territory, when there are health crises that involve serious alterations to normality." The state of alarm was approved by decree on Saturday the 14th of March and, as a consequence, the country was immediately locked down.

bankruptcy, and in March 2020 with the COVID-19 lockdown): on both occasions, real-time updates of the DENSI proved to be very useful in revising downwards the nowcast of economic activity. In a second exercise, we focused on predicting the COVID-19 economic crisis and computed the probability of recession at different horizons from March 2020. Our results showed that the DENSI provides better results compared to the ESI. The probability of recession in March 2020 jumped to about 100% on the 14th March 2020, i.e., the day Spain was locked down. In the second half of the month, the probability of recession reached values between 90% and 100% at any horizon. By contrast, the ESI indicator, which in March 2020 provided a poor signal of the economic situation, yielded a very small probability of recession in the subsequent months.

References

- Alonso, F., J. Ayuso, and J. Mart´ınez Pages´ (1997): "How Informative are Financial Asset Prices in Spain?" Working Papers 9726, Bank of Spain.
- Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," The Quarterly Journal of Economics, 131, 1593–1636.
- Caggiano, G., E. Castelnuovo, and J. M. Figueres (2017): "Economic policy uncertainty and unemployment in the United States: A nonlinear approach," Economics Letters, 151, 31– 34.
- Calomiris, C. W. and H. Mamaysky (2019): "How news and its context drive risk and returns around the world," Journal of Financial Economics, 133, 299 – 336.
- Cervena, M. and M. Schneider ´ (2014): "Short-term forecasting of GDP with a DSGE model augmented by monthly indicators," International Journal of Forecasting, 30, 498–516.
- Colombo, V. (2013): "Economic policy uncertainty in the US: Does it matter for the Euro area," Economics Letters, 121, 39 – 42.
- Combes, S., C. Bortoli, and T. Renault (2018): "Nowcasting GDP Growth by Reading Newspapers," Economics and Statistics, N.505–506, 2018. Big Data and Statistics (Part 1).
- Fontaine, I., L. Didier, and J. Razafindravaosolonirina (2017): "Foreign policy uncertainty shocks and US macroeconomic activity: Evidence from China," Economics Letters, 155, 121–125.
- Fraiberger, S. P. (2016): "News Sentiment and Cross-Country Fluctuations," Tech. rep., available at SSRN: https://ssrn.com/abstract=2730429.

- Gajewski, P. (2014): "Nowcasting Quarterly GDP Dynamics in the Euro Area The Role of Sentiment Indicators," Comparative Economic Research, 17.
- Gayer, C. and B. Marc (2018): "A "New Modesty"? Level Shifts in Survey Data and the Decreasing Trend of "Normal" Growth," European economy - discussion papers 2015, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.
- Ghirelli, C., J. J. Perez, and A. Urtasun ´ (2019): "A new economic policy uncertainty index for Spain," Economics Letters, 182, 64–67.
- Giannone, D., L. Reichlin, and S. Simonelli (2009): "Nowcasting Euro Area Economic Activity in Real-Time: The Role of Confidence Indicators," National Institute Economic Review, 210, 90–97.
- Keeney, M., B. Kennedy, and J. Liebermann (2012): "The value of hard and soft data for short-term forecasting of GDP," Economic Letters 11/EL/12, Central Bank of Ireland.
- Ludvigson, S. C. (2004): "Consumer Confidence and Consumer Spending," Journal of Economic Perspectives, 18, 29–50.
- Mariano, R. S. and Y. Murasawa (2010): "A coincident index, common factors, and monthly real GDP," Oxford Bulletin of Economics and Statistics, 72, 27–46.
- Martínez Serna, M.-I. and E. Navarro (2005): "The Term Structure of Interest Rates and Expected Economic Growth ," Working papers, SSRN.
- Mazzi, G. L., J. Mitchell, and G. Montana (2014): "Density nowcasts and model combination : nowcasting Euro-area GDP growth over the 2008-9 recession," Oxford Bulletin of Economics and Statistics, 76, 233–256.
- Meinen, P. and O. Roehe (2017): "On measuring uncertainty and its impact on investment: Cross-country evidence from the euro area," European Economic Review, 92,

161–179.

- Nyman, R., S. Kapadia, D. Tuckett, D. Gregory, P. Ormerod, and R. Smith (2018): "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment," SSRN Electronic Journal.

- Shapiro, A., M. Sudhof, and D. Wilson (2020): "Measuring News Sentiment," Tech. rep., Federal Reserve Bank of San Francisco, Working Paper 2017–01, https://doi.org/10.24148/wp2017-01.

- Spanish Economic Association (2015): "CF Index of Economic Activity," Spanish Business Cycle Dating Committee.

- Thorsrud, L. A. (2016): "Nowcasting using news topics Big Data versus big bank," Working Paper 20/2016, Norges Bank.

  ——— (2020): "Words are the New Numbers: A Newsy Coincident Index of the Business Cycle," Journal of Business & Economic Statistics, 38, 393–409.

- Wright, J. H. (2006): "The Yield Curve and Predicting Recessions ," Finance and Economics Discussion Series 2006/07, Federal Reserve Board.

Part IV

Learning the Hard Way: Expectations and the U.S. Great Depression

1 Introduction

This research investigates to what extent a light form of bounded rationality can help a simple dynamic general equilibrium (DGE) model to account for the Great Depression of the 1930s in the United States.

Since the pioneering work by Cole and Ohanian (1999), DGE macroeconomics has been used to study the Great Depression, a phenomenon once considered beyond the grasp of equilibrium business cycle theory.[1] The list of contributors is long, and models range from the real-business-cycle (RBC) to the New-Keynesian (NK) variant.[2]

Most of the DGE literature on the Great Depression uses models with rational expectations. While this improves the analytical and numerical tractability of the model, allowing for an easier comparison with the data, it also has some drawback. In particular, rational expectations build on the assumption that the *explanandum* is one occurrence of a regular phenomenon. This is well explained by Lucas:

> 'Insofar as business cycles can be viewed as repeated instances of essentially similar events, it will be reasonable to treat all agents as reacting to cyclical changes as "risk", or to assume their expectations are *rational*, that they have fairly stable arrangements for collecting and processing information, and that they utilize this

With Luca Pensieroso (IRES,UCL)

[1] See De Vroey and Pensieroso (2006)

[2] See Kehoe and Prescott (2007) and the references therein, Bordo et al. (2000), Christiano et al. (2003), Weder (2006), among others. For a survey and an assessment of this literature see Pensieroso (2007), and the debate between Temin (2008) and Kehoe and Prescott (2008).

information in forecasting the future in a stable way, free of systematic and easily correctable bias.' Lucas (1977), p. 224.

If instead the phenomenon at hand is somewhat irregular, that is if it is new or to some extent perceived as unprecedented, then the use of rational expectations is more questionable on theoretical grounds.

Historical accounts suggest the Great Depression is the epitome of an irregular, unexpected event. Romer (1990) provides convincing evidence that uncertainty was radical and pervasive at the onset of the Great Depression. Lucas himself cautions against tackling the Great Depression as an ordinary business cycle, for, he suggests, the Great Depression was a somewhat unique event:[3]

> 'The Great Depression, however, remains a formidable barrier to a completely unbending application of the view that business cycles are all alike.' Lucas (1980), p. 697.

These considerations call for enriching the DGE analysis of the Great Depression with bounded rationality. In this paper, we introduce adaptive learning in the spirit of Evans and Honkapohja (2001), i.e. a light form of bounded rationality, in an otherwise standard RBC model of the Great Depression in the United States. Our model features calibrated shocks to total factor productivity (TFP), taxes on labour and capital and public expenditures. Simulations show that while the model with adaptive learning does slightly better than its rational-expectations counterpart in accounting for the onset of the Great Depression, it clearly outperforms the latter in accounting for its depth and long duration. This suggests that, in the case of big, traumatic events like the Great Depression, uncertainty may be particularly unfavorable to the recovery phase through the expectation channel, a trait by and large overlooked by the literature so far. Notice that in this paper we understand uncertainty in a specific, narrow way. Agents do know the dimension and origin of the shocks, but they

[3] On Lucas's qualms about equilibrium models of the the Great Depression, see De Vroey and Pensieroso (2006)

have a limited understading of how those map into the general equilibrium reaction of the economy. In other words, in our model uncertainty is about the transmission mechanism of the model. Hence, if expectations were rational, there would be no uncertainty according to our definition of it.

The protracted character of the Great Depression has long been considered a puzzling aspect, particularly difficult to account for in a quantitative macroeconomic model. The typical solution to this difficulty has been to complexity the model, by adding an additional set of policy-driven shocks and/or more frictions, on the basis of historical evidence. The most notable contribution in this spirit is Cole and Ohanian (2004), who maintain that the cartelisation and wage policy of the New Deal are major responsables for the delayed recovery from the Depression. Our paper complements their explanation, suggesting that on top of policy shocks, uncertainty might have delayed the recovery significantly through their impact on expectations.

While to our knowledge we are the first to tackle the issue of the Great Depression by means of a DGE model with adaptive learning, we certainly are not the first to suggest a role for expectations in explaining the Great Depression.[4] Limiting to the DGE literature, Eggertsson (2008) argues that Roosevelt's New Deal policies produced a favorable shift in expectations, which explains the acceleration in the growth rate of (undetrended) GDP after 1933; while Harrison and Weder (2006) suggest that sunsposts, i.e. self-fullfilling prophecies unrelated to fundamentals can explain the entire Depression period. Our paper differs from these works in that we i) only assume shocks to the fundamentals and ii) do not model Markov-switching policy regimes, nor do we assume that, absent a policy change, the economy would have slided into the abyss. We introduce, instead, a small deviation from rational expectations and explore in the presence of high uncertainty the consequences of adaptive beliefs in accounting for the Great Depression. In this sense, our work is more akin in method to the literature originated by Cole and Ohanian (1999), Kehoe and Prescott (2002) and

[4]The first explanation of the Great Depression based on expectations is obviously Keynes's *General Theory*, with his insistence on animal spirits. See Keynes (1936).

Prescott (1999), while at the same time stressing the role of uncertainty, under the shape of a light form of bounded rationality. Our approach, in particular, allows us to compare our results to those obtained by Cole and Ohanian (1999) using a standard rational-expectations real-business-cycle (RBC) model. This way, we aim at quantifying the importance of the expectation channel as an additional mechanism to explain the Great Depression. We do not have, however, the pretention of providing a comprehensive explanation of the Great Depression, which would require a more sophisticated model, including a throughout modelisation of opennes to international exchanges and several other real and monetary factors.[5]

The literature on adaptive learning was pioneered by Bray and Savin (1986), Marcet and Sargent (1989) and Evans and Honkapohja (2001). Its gist is to relax the strong assumptions underpinning rational expectations, by relying on the less demanding "cognitive consistency principle": economic agents in the model should be as knowledgeable about the model economy as the economists are about the actual economy. In other words, agents in the model do not necessarily know the full structure of the economy and use forecasting rules to form their expectations. In this sense, they behave as econometricians who attempt to learn the true correlations in the aggregate economy using simple models, whose coefficients are updated every period with the arrival of new information. Therefore, the presence of uncertainty -understood as higher volatility and noisy information- may have an impact on the learning process, shaping the expectation channel in the model.

In models with adaptive learning, the functional form of the forecasting model of the agents – what is known as their 'perceived law of motion' (PLM) – may be specified in different ways. A first way is to assume that agents know the functional form of the model and the exogenous driving processes in the economy, and hence, they are able to find the minimum state variables solving their expectations, as under rational expectations, as shown in Evans and Honkapohja (2001).[6] The key departure here is that, despite of finding the

[5]The economic literature on the Great Depression is too vast and controversial to be conveniently surveyed here. The interested reader may look at Pensieroso and Restout (2021) and the references mentioned ther

[6]This type of learning is further explored in Eusepi and Preston (2011) and Preston (2005). Milani (2007, 2008) study the implications of adaptive learning in medium-scale New Keynesian models

correct functional form, they are uncertain about the true coefficients associated to them and have to learn them. Therefore, instead of plugging in the rational-expectations solution, they estimate the minimal state variable representation on actual data, and use an algorithm to update every period the estimated coefficients as a function of the forecasting error. The forecasting model converges to the rational expectation solution, or a distribution centered around it, depending on the updating algorithm. However, the larger the deviations of the observations from their historical distribution, the larger the deviations in the learning process. This is relevant in the case of the Great Depression, where the size of the shocks hitting the economy and the fiscal response to it increased the uncertainty about the variables behavior, affecting the learning dynamics.

Our approach represents the smallest departure from rational expectations since agents use forecasting rule that consist on a "learnable" solution of the rational expectations equilibria. This kind of learning represents our benchmark model. An alternative approach regards to the model with "misspecified" beliefs. This is the case where agents information is more restricted, making them unable to find the reduced form of the model and the minimum state variables solving their expectations. This implies that there is no longer a learnable mapping between the rational expectation and the learning solution. It represents a larger deviation from the rational expectations hypothesis and it introduces an additional degree of freedom in the choice of the variables to be included in the formation of expectations. As a consequence, in this type of models the PLM is no longer restricted to be consistent with the rational expectations decision rule. One popular approach to this kind of learning is the 'Euler' learning, as in Milani (2011), Slobodyan and Wouters (2012a,b) and Aguilar and Vazquez (2021, 2022), where agents use a set of information limited to the variables appearing in the Euler equation when forming their expectations. In an extension to our benchmark, we build a model of the U.S. Great Depression with a misspecified adaptive learning. Results show that the further we move from rational expectations, the more the sluggish character of the recovery from the Depression is accentuated. This confirms our contention that expectations must have played a significant role in delaying the recovery out of the Great Depression of the 1930s.

The rest of the paper is organized as follows. In Section 2, we present the analytical model, discuss its rational expectations solution and describe learning mechanism considered. In Section 3 we calibrate the model on U.S. data, while simulations are presented and discussed in Section 4. We present an alternative learning algorithm in Section 5 and discuss its implications for the fitting of the model. Finally, Section 6 concludes.

2 The model

2.1 The model economy

The model economy is populated by P representative households and a Government, population is constant. Each household owns one representative firm that uses labour, n, and capital, k, to produce in perfect competition the same indifferentiated good, y, according to:

$$y_t = \exp(z_t) k_{t-1}^{\alpha} \left(x_t n_t\right)^{1-\alpha}, \tag{1}$$

where all variables are expressed in per-capita terms, x is the labour-augmenting technical progress, and z is the stochastic component of total factor productivity (TFP). The former grows at a constant rate, thereby determining the balanced-growth path of the model:

$$x_t = (1+\gamma) x_{t-1}. \tag{2}$$

where all variables are expressed in per-capita terms and z is the stochastic component of total factor productivity (TFP). The stochastic component of TFP is assumed to follow an AR(1) process:

$$z_t = \rho_z z_{t-1} + \varepsilon_t. \tag{3}$$

The representative firm maximizes its profits, which delivers the demand for labour and capital

$$w_t = \exp(z_t)(1-\alpha)k_{t-1}^{\alpha}n_t^{-\alpha}, \tag{4}$$

$$r_t = \exp(z_t)\alpha k_{t-1}^{\alpha-1}n_t^{1-\alpha}, \tag{5}$$

where w is the wage rate and r the rental price of capital.

The household cares for consumption, c, and leisure, $l \equiv (1-n)$. Like in Christiano and Eichenbaum (1992), we assume that consumption is made of services that are related to private and public consumption – c^p and g, respectively – according to

$$c_t = c_t^p + \eta g_t. \tag{6}$$

Though we assume that public expenditures are perceived as exogenous by the household, the Government funds them out of proportional taxation on revenues from labour, τ^w, and capital, τ^r. Both public expenditures and taxes are subject to an AR(1) shock. A lump sum transfer, tr, also perceived as exogenous by the household, ensures that the Government's budget is balanced in every period. Accordingly, the (balanced) budget of the Government reads

$$g_t + tr_t = \tau_t^w w_t n_t + \tau_t^r r_t k_{t-1}, \tag{7}$$

and the shocks linked to the public sector are:

$$g_t = \rho_g g_{t-1} + v_t; \tag{8}$$

$$\tau_t^r = \rho_r \tau_{t-1}^r + u_t; \tag{9}$$

$$\tau_t^w = \rho_w \tau_{t-1}^w + s_t. \tag{10}$$

The household finances consumption and investments, i, out of (net) revenue from labour and capital. Her budget constraint reads

$$c_t^p + i_t = (1 - \tau_t^w) w_t n_t + (1 - \tau_t^r) r_t k_{t-1} + tr_t. \tag{11}$$

Capital accumulates according to

$$k_t = (1 - \delta) k_{t-1} + i_t. \tag{12}$$

For the sake of analytical tractability, we assume a log-additive utility function:

$$U_t = \ln c_t + \varphi \ln (1 - n_t) \tag{13}$$

Given the intertemporal discount factor, β, the representative household chooses (private) consumption, leisure and investment so as to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t U_t, \tag{14}$$

subject to Equations (4), (5), (6), (11), (12) and (13).

The first order conditions of the problem delivers the Euler equation, disciplining the consumption-saving choice,

$$\frac{1}{c_t} = \frac{\beta}{1+\gamma} E_t \left(\frac{1}{c_{t+1}} \Big(\left(1 - \tau_{t+1}^r\right) \left(\alpha \exp(z_{t+1}) \tilde{k}_t^{\alpha-1} x_{t+1} n_{t+1}^{1-\alpha}\right) + 1 - \delta \Big) \right), \tag{15}$$

and the equation determining the equilibrium on the labour market,

$$\frac{\varphi}{1 - n_t} = (1 - \tau_t^w) \frac{(1-\alpha) \exp(z_t) k_{t-1}^{\alpha} x_{t+1} n_t^{1-\alpha}}{c_t}, \tag{16}$$

2.2 Expectations

Modelling expectations explicitly requires to start from the analytical solution of the model. To obtain it, we first linearise the model around the (detrended) steady-state. Then, we solve the linear system, to get the reduced-form expression of the model. As shown in the Appendix, this delivers

$$\hat{k}_t = a_1 E_t \hat{k}_{t+1} + a_2 \hat{k}_{t-1} + b_1 \hat{z}_t + b_2 \hat{g}_t + b_3 \hat{\tau}_t^r + b_4 \hat{\tau}_t^w, \qquad (17)$$

$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \varepsilon_t (17b)$$

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + v_t, (17c)$$

$$\hat{\tau}_t^r = \rho_{\tau^r} \hat{\tau}_{t-1}^r + u_t, (17d)$$

$$\hat{\tau}_t^w = \rho_{\tau^r} \hat{\tau}_{t-1}^w + s_t, (17e)$$

$$\hat{q}_t = \psi_1 \hat{k}_t + \psi_2 \hat{k}_{t-1} + \psi_3 \hat{z}_t + \psi_4 \hat{g}_t + \psi_5 \hat{\tau}_t^r + \psi_6 \hat{\tau}_t^w (17f)$$

with $\hat{q} = \hat{y}, \hat{c}, \hat{n}, \hat{t}r$, and $\hat{y}$ meaning y in log-deviations from steady state and $\psi_1..\psi_6$ are the associated coefficients under the reduced form notation. Hence, given the shocks, the dynamics of capital is the driving force of the system, for the other endogenous variables are known once capital is known. Since the initial conditions are given and known, the model is actually solved once the form of $E_t \hat{k}_{t+1}$ is known. It is at this stage of the argument that the difference between rational expectations and the simplest form of adaptive learning materializes.

2.2.1 Rational expectations

Models with rational expectations are solved using the method of undetermined coefficients (Uhlig (1999)). We first conjecture the solution for capital:

$$\hat{k}_t = \phi_k \hat{k}_{t-1} + \phi_z \hat{z}_t + \phi_g \hat{g}_t + \phi_r \hat{\tau}_t^r + \phi_w \hat{\tau}_t^w, \tag{18}$$

where ϕ_i are the supposed elasticities between capital and the different state variables. The conjecture is simply that capital must depend on its past value and on the shocks. Using Equations (17b), (17c), (17d) and (17e), the conjecture can be written as:

$$\hat{k}_t = \phi_k \hat{k}_{t-1} + \phi_z \rho_z \hat{z}_{t-1} + \phi_g \rho_g \hat{g}_{t-1} + \phi_r \rho_r \hat{\tau}_{t-1}^r + \phi_w \rho_w \hat{\tau}_{t-1}^w + \frac{\phi_z \varepsilon_t}{\rho_z} + \frac{\phi_g v_t}{\rho_g} + \frac{\phi_r u_t}{\rho_r} + \frac{\phi_w s_t}{\rho_w}, \tag{19}$$

Rolling over Equation (19), and noticing that by construction $E_t\left(\varepsilon_{t+1}\right) = 0$, $E_t\left(v_{t+1}\right) = 0$, $E_t\left(u_{t+1}\right) = 0$ and $E_t\left(s_{t+1}\right) = 0$, we obtain:

$$E_t \hat{k}_{t+1} = \phi_k \hat{k}_t + \phi_z \rho_z \hat{z}_t + \phi_g \rho_g \hat{g}_t + \phi_r \rho_r \hat{\tau}_t^r + \phi_w \rho_w \hat{\tau}_t^w. \tag{20}$$

This expression may be plugged in Equation (17). Then, using Equations (17b), (17c), (17d) and (17e), and rearranging terms one gets:

$$\begin{aligned}
\hat{k}_t =& \frac{a_2}{1 - a_1 \phi_k} \hat{k}_{t-1} \\
&+ \frac{(a_1 \phi_z \rho_z + b_1)\rho_z}{1 - a_1 \phi_k} \hat{z}_{t-1} + \frac{(a_1 \phi_g \rho_g + b_2)\rho_g}{1 - a_1 \phi_k} \hat{g}_{t-1} + \frac{(a_1 \phi_r \rho_r + b_3)\rho_r}{1 - a_1 \phi_k} \hat{\tau}_{t-1}^r + \frac{(a_1 \phi_w \rho_w + b_4)\rho_w}{1 - a_1 \phi_k} \hat{\tau}_{t-1}^w \\
&+ \frac{(a_1 \phi_z \rho_z + b_1)}{1 - a_1 \phi_k} \varepsilon_t + \frac{(a_1 \phi_g \rho_g + b_2)}{1 - a_1 \phi_k} v_t + \frac{(a_1 \phi_r \rho_r + b_3)}{1 - a_1 \phi_k} u_t + \frac{(a_1 \phi_w \rho_w + b_4)}{1 - a_1 \phi_k} s_t.
\end{aligned} \tag{21}$$

Finally, since agents have rational expectations, they know their conjecture is true. Accordingly, the coefficients in Equations (21) must coincide with those in Equation (19). Hence, we shall have the following system of equations in ϕ_i:

$$\phi_k = \frac{a_2}{1 - a_1\phi_k}, \qquad \phi_z = \frac{(a_1\phi_z\rho_z + b_1)}{1 - a_1\phi_k},$$

$$\phi_g = \frac{(a_1\phi_g\rho_g + b_2)}{1 - a_1\phi_k}, \qquad \phi_r = \frac{(a_1\phi_r\rho_r + b_3)}{1 - a_1\phi_k},$$

$$\phi_w = \frac{(a_1\phi_w\rho_w + b_4)}{1 - a_1\phi_k},$$

Plugging the solutions of this system into Equation (21), the linear system formed by the latter and Equations (17b), (17c), (17d), (17e) and (17f) gives the rational expectations solution to the reduced-form model.[7]

2.2.2 Adaptive learning

Introducing adaptive learning instead of rational expectations is relatively straightforward in this class of models. Following Evans and Honkapohja (2001), we start from Equation (20) and simply assume that agents cannot know if, and to what extent, their conjecture is true. Therefore, they cannot equate Equations (21) and Equation (19) and solve for the ϕ_i coefficients. Instead, they have to estimate them on actual data, like an econometrician would do. Hence, the ϕ_i coefficients will now be time-dependent: period-by-period, agents will update their estimations, taking the estimation error into account, thereby generating a process of dynamic adaptive learning. Accordingly, the conjecture under adaptive learning, also known as 'perceived law of motion' (PLM) will be

$$E_t\hat{k}_{t+1} = \phi_{k,t-1}\hat{k}_t + \phi_{z,t-1}\hat{z}_t + \phi_{g,t-1}\hat{g}_t + \phi_{r,t-1}\hat{\tau}_t^r + \phi_{w,t-1}\hat{\tau}_t^w. \tag{22}$$

[7] The system admits a unique stationary solution if and only if $|a_1\phi_k + a_2| < 1$ (see Evans and Honkapohja (2001)).

This is basically Equation(20) with the expectation coefficients ϕ_i now indexed at $t-1$. The reason for this indexation is that i) the coefficients are time-dependent and ii) they are updated after every realization of k, which is a driving force of the system in addition to the exogenous shock processes. Since k is a state variable, this means that the coefficients estimated in t are conditional to the information set in $t-1$.

Following the same procedure as in the rational expectations case, conjecture (22) may be plugged in Equation (17). Then, using Equations (17b), (17c), (17d), and (17e), and rearranging terms one gets the so-called 'actual law of motion' (ALM):

$$\begin{aligned}\hat{k}_t &= \frac{a_2}{1-a_1\phi_{k,t-1}}\hat{k}_{t-1} + \frac{(a_1\phi_{z,t-1}\rho_z + b_1)\rho_z}{1-a_1\phi_{k,t-1}}\hat{z}_{t-1} \\ &+ \frac{(a_1\phi_{g,t-1}\rho_g + b_2)\rho_g}{1-a_1\phi_{k,t-1}}\hat{g}_{t-1} + \frac{(a_1\phi_{r,t-1}\rho_r + b_3)\rho_r}{1-a_1\phi_{k,t-1}}\hat{\tau}^r_{t-1} + \frac{(a_1\phi_{w,t-1}\rho_w + b_4)\rho_w}{1-a_1\phi_{k,t-1}}\hat{\tau}^w_{t-1} \\ &+ \frac{(a_1\phi_{z,t-1}\rho_z + b_1)}{1-a_1\phi_{k,t-1}}\varepsilon_t + \frac{(a_1\phi_{g,t-1}\rho_g + b_2)}{1-a_1\phi_{k,t-1}}v_t + \frac{(a_1\phi_{r,t-1}\rho_r + b_3)}{1-a_1\phi_{k,t-1}}u_t + \frac{(a_1\phi_{w,t-1}\rho_w + b_4)}{1-a_1\phi_{k,t-1}}s_t\end{aligned} \tag{23}$$

The linear system formed by the ALM – Equation (23) – and Equations(17b), (17c), (17d), (17e) and (17f) gives the adaptive learning solution to the reduced-form model. Like in the case of rational expectations, this system has a unique stationary solution if and only if $|a_1\phi_{k,t-1} + a_2| < 1$ (see Evans and Honkapohja (2001)).

To complete the adaptive learning model, it remains now to specify the updating algorithm for the expectation coefficients ϕ_i. Following Evans and Honkapohja (2001) and Marcet and Sargent (1989), we assume:

$$\mathbf{\Phi_t} = \mathbf{\Phi_{t-1}} + \zeta_t \mathbf{R_{t-1}^{-1}} \mathbf{x_{t-1}} (\hat{k}_t - \mathbf{x'_{t-1}} \mathbf{\Phi_{t-1}}), \tag{24}$$

$$\mathbf{R_t} = \mathbf{R_{t-1}} + \zeta_t \big(\mathbf{x_{t-1}} \mathbf{x'_{t-1}} - \mathbf{R_{t-1}}\big), \tag{25}$$

where $\mathbf{x_t}$ is the (column) vector of the state variables $(\hat{k}_t,\ \hat{z}_t,\ \hat{g}_t,\ \hat{\tau}^r_t,\ \hat{\tau}^w_t)$, $\mathbf{R_t}$ is the matrix (of order 5) of the variances and covariances of the state variables and $\mathbf{\Phi_t}$ is the (column) vector of the expectation coefficients (ϕ_k, ϕ_z, ϕ_g, ϕ_r, ϕ_w). The scalar ζ_t stands for the speed of learning. It may be fixed, i.e. $\zeta_t = \zeta$ for any t, in which case the algorithm is known as 'constant gain' (CG) learning; or it may be time-dependent, i.e. $\zeta_t = 1/t$, in which case the algorithm is known as 'recursive least squares' (RLS) learning. The main difference between these two algorithms regards the speed of updating (gain) and the asymptotic properties.[8] The RLS algorithm updates according to a decreasing sequence, and it can be shown to converge to the rational expectations solution, provided the latter is unique and stationary. The CG algorithm updates according to a constant sequence, and converge to a distribution centered around the rational expectations solution. This introduces perpetual learning in the model. As the system indicates, every period the set of coefficients is updated according to the forecast error, $(\hat{k}_t - \mathbf{x'_{t-1}\Phi_{t-1}})$, corrected by the variance of the state variables, $\mathbf{R_{t-1}^{-1}x_{t-1}}$, and the speed of learning, ζ_t. The variance-covariance matrix ($\mathbf{R_t}$) is also updated every period taking into account the variation of the volatility and cross-correlation of the state variables, corrected again by the speed of learning, ζ_t.[9]

3 Calibration

The model's structural parameters are calibrated as shown in Table 1. The unit period is the year. As customary in the literature, we assume that 1929 in the data corresponds to the steady state in the model. The depreciation rate of capital, δ, the labour share in the production function, α, the average growth rate of the U.S. GDP per capita, γ, and the

[8]See Carceles-Poveda and Giannitsarou (2007) and Evans and Honkapohja (2001), for an in depth analysis of the convergence properties of the two algorithms.

[9]To help the reader understanding the updating algorithm, consider the case in which public expenditures and taxes were constantly equal to zero. In this case the system (24)–(25) would take the following form:

$$\begin{pmatrix} \phi_{k,t} \\ \phi_{z,t} \end{pmatrix} = \begin{pmatrix} \phi_{k,t-1} \\ \phi_{z,t-1} \end{pmatrix} + \zeta_t \begin{pmatrix} \sigma^2_{k,t-1} & \sigma_{kz,t-1} \\ \sigma_{zk,t-1} & \sigma^2_{z,t-1} \end{pmatrix}^{-1} \left[\hat{k}_t - \begin{pmatrix} \hat{k}_{t-1} & \hat{z}_{t-1} \end{pmatrix} \begin{pmatrix} \phi_{k,t-1} \\ \phi_{z,t-1} \end{pmatrix} \right]$$

$$\begin{pmatrix} \sigma^2_{k,t} & \sigma_{kz,t} \\ \sigma_{zk,t} & \sigma^2_{z,t} \end{pmatrix} = \begin{pmatrix} \sigma^2_{k,t-1} & \sigma_{kz,t-1} \\ \sigma_{zk,t-1} & \sigma^2_{z,t-1} \end{pmatrix} + \zeta_t \left[\begin{pmatrix} \hat{k}_{t-1} \\ \hat{z}_{t-1} \end{pmatrix} \begin{pmatrix} \hat{k}_{t-1} & \hat{z}_{t-1} \end{pmatrix} - \begin{pmatrix} \sigma^2_{k,t-1} & \sigma_{kz,t-1} \\ \sigma_{zk,t-1} & \sigma^2_{z,t-1} \end{pmatrix} \right]$$

intertemporal discount factor, β, are fixed as in Cole and Ohanian (1999). The implied steady-state real interest rate, net of taxes and the depreciation rate, is equal to 6.8%. The preference for leisure, φ, is calibrated so that in equilibrium hours worked are 1/3 of the household's time endowment, which is normalized to 1. We calibrate the initial value of public expenditure so that the ratio of public expenditure over GDP in steady-state matches the 1929 value, i.e. 13%. This is not far from the average g/y observed in the period 1919-1929, which is 11% according to data from Cole and Ohanian (1999. The steady-state values for the tax on wages and capital are fixed to the 1929 actual value reported by Joines (1981), i.e., 3.3% and 20% respectively. We assume that public expenditures are perfect substitute for private consumption in the utility function, by setting $\eta = 1$. We have checked that results do not change appreciably if we set $\eta = 0$ instead.

Table 1. Calibrated parameters

Parameters	Value	Parameters	Value
γ	0.019	τ^{wss}	0.035
φ	1.639	τ^{rss}	0.20
β	0.972	ρ_g	0.678
δ	0.10	ρ_z	0.847
α	0.33	ρ_r	0
g_{ss}	0.058	ρ_w	0
η	1	ζ	0.03

TFP shocks are calculated as residuals from regressing an AR(1) process on detrended TFP data from Cole and Ohanian (1999, in line with Equation (3). The autoregressive coefficient ρ_z for this period is estimated to 0.85.[10] Similarly, public expenditure shocks are the residuals from regressing an AR(1) process on the detrended share of public expenditure

[10]Notice that while we use Equation 3 to estimate the Solow residual, the model simulation of the production function does not necessarily match the observed data, due to the general equilibrium response of the model to the calibrated shocks. This is a general property of RBC models and holds true under both adaptive learning and rational expectations.

over GDP, which reduces to Equation (8). The autoregressive coefficient ρ_g is estimated to 0.68. Tax shocks are directly measured from the data as the difference between the observed value at each time t and their 1929 (steady-state) value. Accordingly, we set $\rho_w = \rho_r = 0$.

The initial conditions for the learning model

In the case of the model with adaptive learning, we need to give initial conditions for equations (24) and (25), and a value to the gain parameter ζ. We set $\zeta = 0.03$, in line with the accepted range in the literature, which is between 0.02 and 0.06.[11] For what concerns the ϕ_i coefficients, we assume them to start at their rational expectations value. For what concerns the variance and covariance matrix, $\mathbf{R}$, we use the one implied by the rational-expectations model, obtained by simulating the model for 250 periods under random shocks. For this exercise, the volatility of the shocks is chosen to match the observed volatility of output in the periods 1919-1929 and 1948-1975.[12] Additionally in the appendix we show that the results are robust to the simulations under a calibration with a 50% higher and lower volatility.

4 Simulations

The model period is the year. All variables are assumed to be at steady state in 1929. We feed in the calibrated values for the shocks on TFP, taxes and public expenditures and run simulations. In Figure 1, we plot the dynamic behavior of the model economy with rational expectations (black-dotted line) and adaptive learning (blue line with markers) against the detrended data (black continuous line). The graphs show that the model with learning behaves similarly to the one with rational expectations for the period 1929-1932. The 1933 trough is deeper in the model with learning, which accounts for 63% of the cumulative drop

[11] In Appendix 4 we discuss the sensibility of our results to different values of ζ, including the recursive-least-square formulation.

[12] This implies a simulation with a standard deviation of 3% for the TFP shock, 2.68% for the public expenditure shock and 5.4% and 3.15% for the tax shock on capital and labour income, respectively. These numbers are consistent with the observed volatility of TFP, public expenditure and taxes on capital and labour income in that period, namely 2.11%, 2.25%, 3.12% and 2.82% respectively.

in detrended output, compared to 57% for the model with rational expectations.

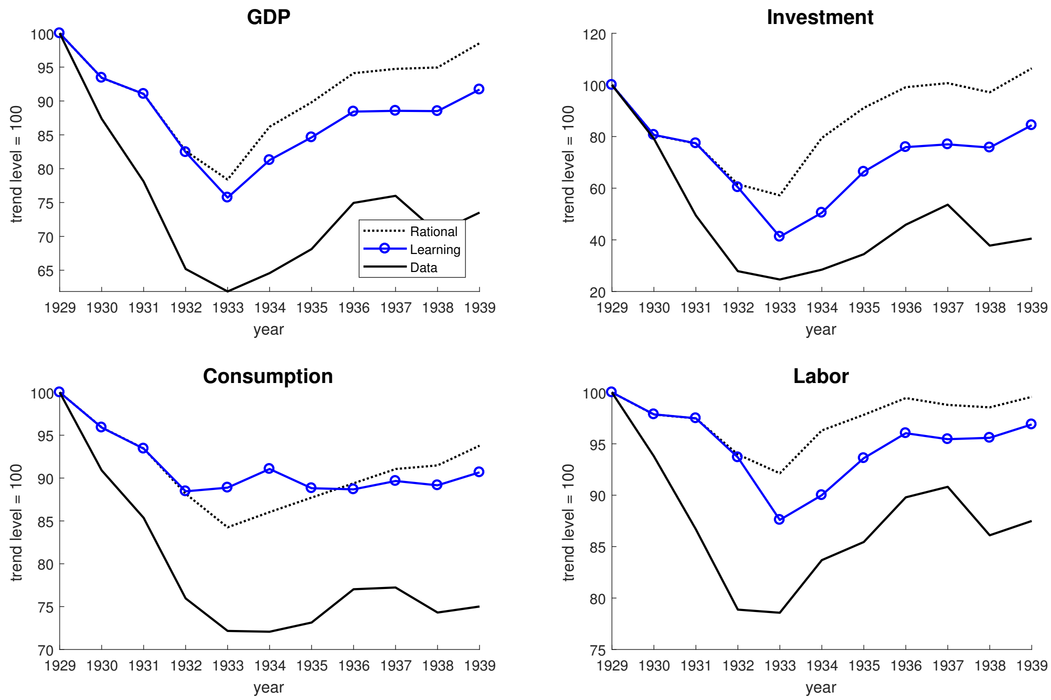


Figure 1: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with markers), data (black line).

The greatest difference between the two models, however, is observed in the 1933-1939 period. During this period all variables, with the exception of consumption, are significantly lower in the model with learning than in the model with rational expectations. By 1939, GDP is almost at trend in the model with rational expectations, about 9 points below trend in the model with adaptive learning, and about 26 points below trend in the data.[13] Thus, the model with learning accounts for 32% of the cumulative drop in detrended output in 1939, compared with a meager 6% for the rational expectations model. These results lead to two conclusions. First, bounded rationality helps a simple DGE model to account for the Great Depression of the 1930s, especially for its depth and long duration. Second, expectations are not likely to be the whole story. The model still needs exogenous shocks to the fundamentals (real shocks,

[13]As shown in Appendix 4, the results mentioned in the text are independent of the initialization of the variance-covariance matrix **R**. We used random generated data as benchmark, but the model with theoretical moments does just as good.

in the case of our simple, flexible price model). Bounded rationality in the form of adaptive learning acts as an amplifier for the shocks. This conforms to the historical literature that sees expectations as an element of disturbance, hampering the business environment of the 1930s, deepening the recession and delaying the recovery. It is also in line with the standpoint of Cole and Ohanian (2004), who claim that additional shocks are needed on top of standard monetary and real shocks if one is to explain the long duration of the Great Depression.

In order to grasp the working of the model, it may be useful to refer to Figure 2. There, we reproduce two sets of graphs in two columns. In the first column to the left, we have the simulated pattern of ϕ_i. In the second column, the deviations from steady state of all the state variables. Hence, this figure represents a visual decomposition of the ALM – Equation (21) or (23), with rational or adaptive expectations, respectively – element by element.

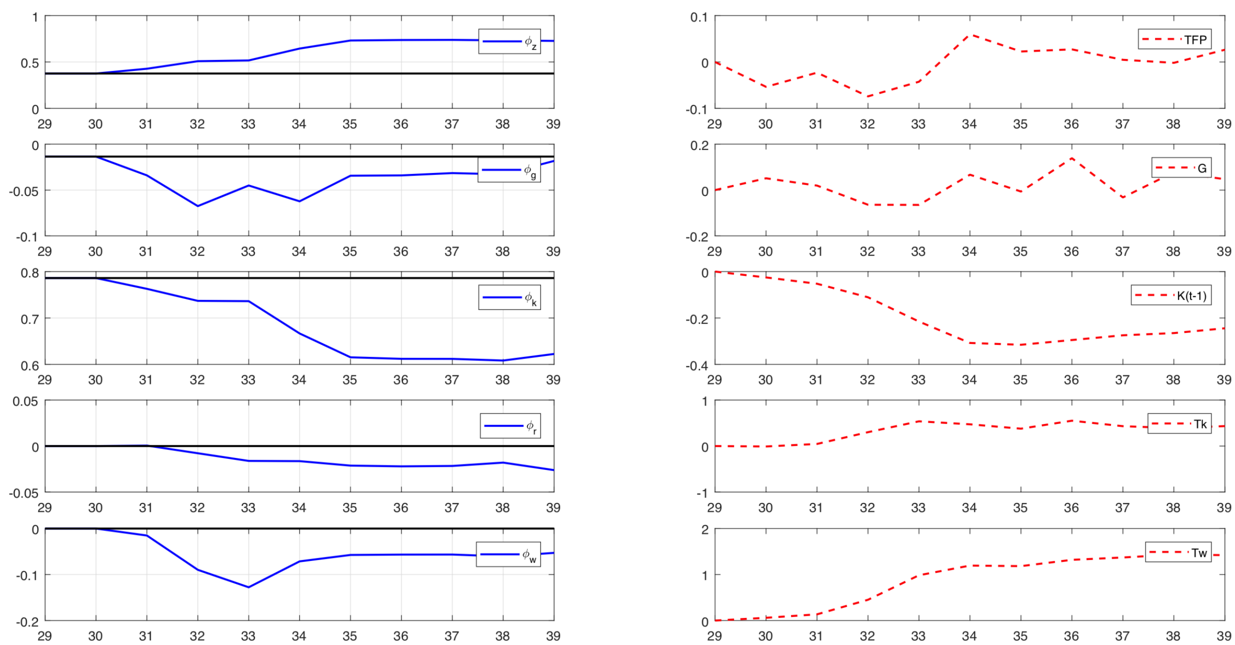


Figure 2: Simulations. Expectation coefficients (left column) and shocks (right column). Blue line: adaptive learning. Black line: rational expectations. Red line: data.

The graphs show that while with rational expectations the expectation coefficients are obviously constant, they do vary a lot in the model with adaptive learning. In particular, the elasticity of the expected value of capital to its past value decreases significantly along the Depression. At the same time, the elasticities of the expected value of capital to tax and public expenditures shocks become negative, while that to TFP shocks becomes positive. Accordingly, the increase in taxes and public expenditures in the second part of the 1930s have a strong negative impact on the expected value of capital, while the model reacts less to the negative dynamics of capital in the past. These effects contribute to further depress the economy, and contrast the tendency to recovery driven by the positive TFP shocks since 1934.

The economic intuition behind the working of the model can be resumed as follows. With rational expectations, the dynamics of the model is fully determined by the persistence of the shocks and the dynamics of capital. In the model with learning, instead, the after-shock dynamics are also driven by revisions in beliefs, which operate through changes in the agent's estimation of the elasticity of capital with respect to its past values and contemporary shocks. Since simulations in both models start from the rational-expectations solution, it takes some time before changes in beliefs matter. The two models, hence, overlap at the onset of the Depression. As soon as agents adapt their expectations, the capital-labour ratio drops dramatically in the learning model, and remains constantly and significantly below its rational-expectations counterpart (see Figure 4 in Appendix). The implied higher value of the real interest rate delays the recovery. This pattern is accentuated by the additional impact on expectations brought about by fiscal policy. As shown in Figure 2, the tax shocks turn out to be progressively more important, while their impact on the dynamics of capital through expectations (the ϕ_i), which would be zero under rational expectations, becomes more and more negative.

In fact, the presence of several shocks is crucial to the results. Since simulations start from the rational expectations solution, beliefs only matter if the economic environment is sufficiently noisy[14]. In a model with TFP shocks only, for instance, the type of learning we have been discussing so far would have little impact on the dynamics of aggregate variables.[15]

[14] In the appendix we show that results are robust to starting the learning in the early 20s

[15] Results available upon request. Notice that there exist models in which TFP-news shocks can cause

When the economy is perturbed along more than one dimension, instead, forecast errors add to each other, thereby amplifying the distorting role of beliefs. The reasonale behind is that agents face a period with higher volatility and forecasting errors than the historically observed. With few realizations agents are uncertain on whether this is temporary or some of the variables distribution has now changed. This is relevant in the case of public expenditure and taxes, whose presence in the economy increased since the Great Depression. This illustrates that uncertainty about expectations matters more in complex, deep crises.[16]

5 Extensions

The learning solution implemented here above is but one of the possible ways of introducing learning in DSGE models. We have chosen it since it witnesses the smallest departure from rational expectations. This has allowed us to assess the impact of bounded rationality on the data mimicking ability of the model for the Great Depression of the 1930s, in the case that is most unfavorable to it, i.e. in the case in which the departure from rational expectations is only minimal. We are now going to study how modelling learning in different ways impacts on the explanation of the Great Depression. The general lesson we may draw from this exercise is twofold. First, the main conclusion from our previous analysis holds true, whatever the type of learning considered: uncertainty helps to explain in particular the slow recovery from the Great Depression. Second, the further the model departs from rational expectations, the stronger the Depression is.

While modelling adaptive learning in Section 2.2.2, we have assumed that agents know

macroeconomic fluctuations of important magnitude, like for instance Beaudry and Portier (2004) or Jaimovich and Rebelo (2009), see Beaudry and Portier (2014) for a survey. This is yet another, different way in which expectations may drive the business cycle. It differs from our approach, since we do not model news and have bounded rationality, while they typically assume rational expectations. Introducing adaptive learning in this kind of models and exploring their predictive power for the Great Depression is in interesting direction for future research.

[16] This is due to learning updating scheme. Several large shocks induce more volatility and higher correlations in the variance-covariance matrix Rt associated with the learning process. Hence, even though the shocks are orthogonal, and the model converges to the RE equilibrium in the long run, still in the short run the variance-covariance matrix may have non-zero cross-correlations, when volatility is high and the number of observables relatively low.

the full structure of the model and are able to obtain its reduced-form solution. Following Milani (2007) and Slobodyan and Wouters (2012a,b), we now relax this assumption and assume instead that agents can only derive the model up the first order conditions. This implies they have to solve the expectation operator appearing in the Euler equation (15) without the possibility of iterating terms. Such a solution stems from models with misspecified beliefs, in which agents have to solve their expectations using a rather limited set of information compared to the minimum state variable approach.[17] While the assumption of limited information may be more realistic one technical implication of this approach is that there is no longer a learnable mapping between the rational expectation and the learning solution.[18]

In order to minimize the departure from rational expectations, we assume that households, yet don't know the reduce form of the model, expect future consumption and hours worked to be function of the state variables. Hence, the perceived law of motion will be

$$E_t\hat{c}_{t+1} = \phi^c_{k,\,t-1}\hat{k}_t + \phi^c_{z,\,t-1}\hat{z}_t + \phi^c_{g,\,t-1}\hat{g}_t + \phi^c_{\tau^r,\,t-1}\hat{\tau}^r_t + \phi^c_{\tau^w,\,t-1}\hat{\tau}^w_t. \tag{26}$$

$$E_t\hat{n}_{t+1} = \phi^n_{k,\,t-1}\hat{k}_t + \phi^n_{z,\,t-1}\hat{z}_t + \phi^n_{g,\,t-1}\hat{g}_t + \phi^n_{\tau^r,\,t-1}\hat{\tau}^r_t + \phi^n_{\tau^w,\,t-1}\hat{\tau}^w_t, \tag{27}$$

where ϕ^c_x and ϕ^n_x are the estimated coefficients associated with the PLM for the expectations of consumption and hours worked respectively for each state variable x. As clear from the PLMs, the parameter coefficients are no longer a combination of the reduced form parameters, which explains why there is no longer a mapping between the solution in in this set up and the rational expectations solution. As a result, this approach amplifies the limited information issue that agents are facing, which in turns amplifies the overall effect of expectations on the economy

[17]Slobodyan and Wouters (2012a,b) argue that assuming that agents know the reduced form of the model when estimating the correct minimum state variable representation of the model in their expectations, specially in large models such as theirs, is an extreme assumption.

[18]As shown in Appendix D, while he model with this type of learning amplifies the volatility of economic variables and increase the persistence of fluctuations, it still maintains the same average behaviour over time as the model with rational expectations.

The updating algorithm for the PLM coefficients is similar to Equations (24) and (25):

$$\mathbf{\Phi_t} = \mathbf{\Phi_{t-1}} + \zeta_t \mathbf{R_{t-1}^{-1}} \mathbf{x_{t-1}} (\mathbf{\hat{f}_t} - \mathbf{x'_{t-1}} \mathbf{\Phi_{t-1}}), \tag{28}$$

$$\mathbf{R_t} = \mathbf{R_{t-1}} + \zeta_t \big(\mathbf{x_{t-1}} \mathbf{x'_{t-1}} - \mathbf{R_{t-1}} \big), \tag{29}$$

where $\mathbf{\hat{f}}$ is the (column) vector of the forecasted variables $\hat{c}$ and $\hat{n}$, $\mathbf{x_t}$ is the (column) vector of the state variables ($\hat{k}$, $\hat{z}$, $\hat{g}$, $\hat{\tau}^r$, $\hat{\tau}^w$), $\mathbf{R_t}$ is the matrix (of order 5) of the variances and covariances of the state variables, $\mathbf{\Phi_t}$ is the matrix of the coefficients corresponding to the PLMs, and ζ_t stands for the speed of learning.

We initialize the model by running simulations with random generated data using the same set of shocks as in Section 3, which also allows us to calculate the initial conditions for the ϕ_i^j. We then plug in the measured shocks linked to the Great Depression and simulate the model.

Figure 3 compares the dynamics of the model with learning about FOCs ('Euler learning' – red line with markers) with that of the model with rational expectations (black-dotted line) and adaptive learning (blue line with markers), and contrasts them with the detrended data (black continues line).

The graphs shows that while learning about FOCs delivers a dynamics that is qualitatively similar to the benchmark adaptive learning model, it greatly amplifies the response of the model to the shocks, so that the model with learning about FOCs tracks the data on output, investment and employment much better than the benchmark. This confirms that the farer we depart from rational expectations, the more a simple DGE model can account for the the Great Depression of the 1930s.

To better understand the logic of results, remember that in this model the after-shock dynamics are driven by revisions in beliefs. The importance of this factor is even stronger in this extension than in the benchmark model, because here expectations enter directly the consumption and labour equations, while in the benchmark they only did indirectly through the law of motion of capital. In this learning case, agents will expect a stronger drop in labour and a higher capital-labour ratio (see Figure 4 in Appendix B), which coupled with the actual and expected lower drop in real wages, wich results in a higher intertemporal substitution effect, with investment decreasing more than in the rational-expectations and the baseline adaptive-learning cases, and consumption, instead, increasing. As the drop in investment progressively affects capital, the capital-labour ratio eventually diminishes and remains significantly below its counterpart in both the rational-expectations and the baseline adaptive models. The implied higher value of the real interest rate delays the recovery further. So, under this approach, learning affects the onset of the Great Depression, and has an even stronger impact on its duration.

Notice that the model overestimates the dynamics of consumption in the midst of the 1930s. This is due to the simplified, supply-driven nature of the model, coupled with the abnormal magnitude of the shocks. With investment dropping abruptly and public expenditures following an exogenous given pattern, consumption is the only variable of adjustment to maintain the equilibrium between aggregate demand and aggregate supply.[19] Additionally,

In Appendix 4, we explore alternative modelling of learning expectations: hybrid expectations, learning with a constant. These variants deliver results that are in line with those of Section 4, both qualitatively and quantitatively.

[19]One could improve the performance of the model with respect to consumption by imposing a higher relative risk aversion and a higher elasticity of labour supply with respect to wages, while at the same time severing the link between labour supply and the marginal utility of consumption. In a robustness exercise available upon request, we have simulated the same model with GHH preferences (Greenwood et al. (1988)), which assumes away the wealth effect on labour supply. We found that high values of both the relative risk aversion (≥ 2.5) and the Frisch elasticity of substitution (~ 2) bring the pattern of consumption more in line with the data. The reason is twofold. First, labour reacts more to the decrease in real wages implied by the shocks. Second, the lower elasticity of intertemporal substitution induces a lesser decrease in investment with respect to the benchmark utility function. Overall, however, the improvement in the dynamics of consumption is obtained at the price of decreasing the performance of the model in terms of output and investment. Alternatively, one may switch to a more complex New-Keynesian model with habit persistence, price and wage stickiness, which would make the model dynamics more demand-driven. Since our focus here is on the role of expectations on output dynamics more than on the perfect data mimicking of its sectoral components, we abstract from all these complications.

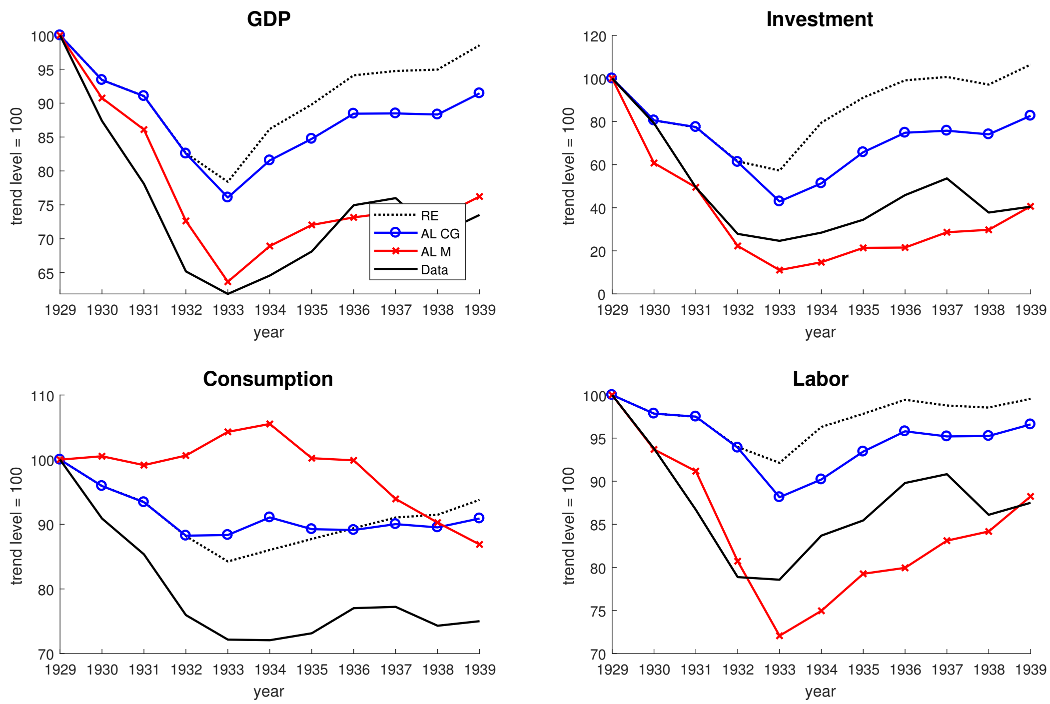


Figure 3: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with circles), model with learning about FOCs (red line with crosses), data (black line).

6 Conclusions

This article introduces adaptive learning in a stylized DGE model of the U.S. Great Depression of the 1930s, thereby contributing to assessing the quantitative importance of uncertainty in agents expectations as a driving factor of the Depression.

Results from simulations show that the the calibrated model with adaptive learning outperforms its rational-expectations counterpart, especially for what concerns the explanation of the slow recovery from the Depression. By 1939, the model with adaptive learning can account for 32% of the actual drop in detrended GDP, as opposed to 6% when expectations are fully rational. This points to uncertainty as an important factor delaying the recovery after 1933. Results are robust to different specifications of the learning dynamics, while the data mimicking ability increases as the model gets further away from rational expectations.

Our analysis leads to conclude that considering uncertainty in the form of a light departure from rational expectations is a promising way to take its important role in big crises into account, without introducing animal spirits or forsaking the analytical advantage and quantitative strength of DGE modelling.

In this article, we have willingly considered the baseline DGE model only, with homogeneous agents, perfect competition and flexible prices. This has allowed us to study the effect of uncertainty on macroeconomic aggregates in the case in which there is no other demand factor driving economic fluctuations. While this may be apposite to our purpose of isolating and quantifying the potential role of uncertainty in the 1930s, we aknowledge that a full-fledged explanation of the Great Depression cannot possibly ignore them.

References

- Aguilar, P., and J. Vázquez, "An estimated DSGE model with learning based on term structure information," Macroeconomic Dynamics, 2021, 1–31.

- Beaudry, P. and F. Portier, (2004). "An exploration into Pigou's theory of cycles," Journal of Monetary Economics, 51(6), 1183-1216

- Bernanke, Ben S., Michael T. Kiley, and John M. Roberts, "Monetary Policy Strategies for a Low-Rate Environment," AEA Papers and Proceedings, May 2019, 109, 421–26.

- Bordo, Michael D., Christopher J. Erceg, and Charles L. Evans, "Money, Sticky Wages and the Great Depression," American Economic Review, 2000, 90, 1447–1463.

- Bray, Margaret M and Nathan E Savin, "Rational expectations equilibria, learning, and model specification," Econometrica: Journal of the Econometric Society, 1986, pp. 1129–1160. 19

- Carceles-Poveda, Eva and Chryssi Giannitsarou, "Adaptive learning in practice," Journal of Economic Dynamics and Control, 2007, 31 (8), 2659– 2697.

- Christiano, L. and M. Eichenbaum, "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," American Economic Review, 1992, 82, 430–450.

  – , R. Motto, and M. Rostagno, "The Great Depression and the FriedmanSchwartz Hypothesis," Journal of Money, Credit and Banking, 2003, 35 (6), 1119–1197.

- Cole, H. L. and L. E. Ohanian, "The Great Depression in the United States from a Neoclassical Perspective," Federal Reserve of Minneapolis Quarterly Review, 1999, 23, 2–24. and ,

  – "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis," Journal of Political Economy, 2004, 112, 779–816.

- De Vroey, Michel and Luca Pensieroso, “Real Business Cycle Theory and the Great Depression: the Abandonement of the Abstentionist Viewpoint,” Contributions to Macroeconomics, 2006, 6, issue 1, article 13.

- Eggertsson, G. B., “Great Expectations and the End of the Depression,” American Economic Review, 2008, 98, 1476–1516.

- Eusepi, Stefano and Bruce Preston, “Expectations, learning, and business cycle fluctuations,” American Economic Review, 2011, 101 (6), 2844–72.

- Evans, G. and S. Honkapohja, Learning and Expectations in Economics, Princeton University Press, Princeton, 2001.

- Gertler, M., “Rethinking the Power of Forward Guidance: Lessons from Japan,” Monetary and Economic Studies, 2017.

- Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman, “Investment, Capacity Utilization, and the Real Business Cycle,” The American Economic Review, 1988, 78 (3), 402–417.

- Harrison, S. G. and M. Weder, “Did Sunspot Forces Cause the Great Depression?,” Journal of Monetary Economics, 2006, 53, 1327–1339.

- Jaimovich, N., and S. Rebelo, (2009). “Can News about the Future Drive the Business Cycle?,” American Economic Review, 99(4), 1097-1118.

- Joines, D. H., “Estimates of Effective Marginal Tax Rates on Factor Incomes,” Journal of Business, 1981, 54, 191–226.

- Kehoe, T. J. and E. C. Prescott, “Great Depressions of the 20th Century,” Review of Economic Dynamics, 2002, 5, 1–18. and ,

  - Great depressions of the twentieth century, Federal Reserve Bank of Minneapolis, 2007. and , “Using the General Equilibrium Growth Model to Study Great

Depressions: a Reply to Temin," Research Department Staff Report 418, Federal Reserve Bank of Minneapolis 2008.

- Keynes, J. M., The General Theory of Employment, Interest, and Money, Macmillan, London, 1936.

- Lucas, R. E., "Understanding Business Cycles," in "Studies in Business Cycle Theory," MIT Press, Cambridge, USA, 1981, 1977, pp. 215–239. ,

  – "Methods and Problems in Business Cycle Theory," in "Studies in Business Cycle Theory," MIT Press, Cambridge, USA, 1981, 1980, pp. 271–296.

- Marcet, Albert and Thomas J Sargent, "Convergence of least squares learning mechanisms in self-referential linear stochastic models," Journal of Economic theory, 1989, 48 (2), 337–368.

- Milani, Fabio, "Expectations, learning and macroeconomic persistence," Journal of monetary Economics, 2007, 54 (7), 2065–2082. ,

  – "Learning, monetary policy rules, and macroeconomic stability," Journal of Economic Dynamics and Control, 2008, 32 (10), 3148–3165. ,

  – "Expectation shocks and learning as drivers of the business cycle," The Economic Journal, 2011, 121 (552), 379–401.

- Pensieroso, L., "Real Business Cycle Models of the Great Depression: A Critical Survey," Journal of Economic Surveys, 2007, 21, 110–142.

- Prescott, E. C., "Some Observations on the Great Depression," Federal Reserve Bank of Minneapolis Quarterly Review, 1999, 23, 25–31.

- Preston, B., "Learning About Monetary Policy Rules When Long-Horizon Expectations Matter.," International Journal of Central Banking, 2005, 1, 81–126.

- Romer, C., “The Great Crush and the Onset of the Great Depression,” Quarterly Journal of Economics, 1990, 105, 597–624. 21
- Slobodyan, S, and R. Wouters, “Learning in a medium-scale DSGE model with expectations based on small forecasting models,” American Economic Journal: Macroeconomics, 2012, 4 (2), 65–101. and ,
  - “Learning in an estimated medium-scale DSGE model,” Journal of Economic Dynamics and control, 2012, 36 (1), 26–46.
- Temin, P., “Real Business Cycle Views of the Great Depression and Recent Events: a Review of Timothy J. Kehoe and Edward C. Prescott’s Great Depressions of the Twentieth Century,” Journal of Economic Literature, 2008, 46, 669–684.
- Uhlig, H., “A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily,” in R. Marimon and A. Scott, eds., Computational Methods for the Study of Dynamic Economies, Oxford University Press, Oxford, 1999, pp. 30–61.
- Vázquez J., and P. Aguilar, “An estimated DSGE model with learning based on term structure information,” Macroeconomic Dynamics, 2021, pp. 1–31. and , “Adaptive learning with term structure information,” European Economic Review, 2021, 134, 103689.
- Weder, M., “The Role of Preference Shocks and Capital Utilization in the Great Depression,” International Economic Review, 2006, 47, 1247–1268.

Appendix Chapter 1

Supplementary appendix Part I

Log-linearized dynamic equations

In addition to equations (1) with $j = 1$ and (3) with $j = 4$ characterizing the 1-quarter rate and the 1-year bond yield, respectively, the set of the remaining log-linearized dynamic equations are the following:

- Aggregate resource constraint:

$$y_t = c_y c_t + i_y i_t + z_y z_t + \varepsilon_t^g, \tag{1}$$

  where $c_y = \frac{C}{Y} = 1 - g_y - i_y$, $i_y = \frac{I}{Y} = (\gamma - 1 + \delta)\frac{K}{Y}$, and $z_y = r^k \frac{K}{Y}$ are steady-state ratios. As in Smets and Wouters (2007), the depreciation rate and the exogenous spending-GDP ratio are fixed in the estimation procedure at $\delta = 0.025$ and $g_y = 0.18$.

- Investment equation:

$$i_t = i_1 i_{t-1} + (1 - i_1) E_t i_{t+1} + i_2 q_t + \varepsilon_t^i, \tag{2}$$

  where $i_1 = \frac{1}{1+\overline{\beta}}$, and $i_2 = \frac{1}{(1+\overline{\beta})\gamma^2\varphi}$ with $\overline{\beta} = \beta\gamma^{(1-\sigma_c)}$.

- Arbitrage condition (value of capital, q_t):

$$q_t = q_1 E_t q_{t+1} + (1 - q_1) E_t r_{t+1}^k - (r_t - E_t \pi_{t+1}) + c_3^{-1} \varepsilon_t^b, \tag{3}$$

  where $q_1 = \overline{\beta}\gamma^{-1}(1-\delta) = \frac{(1-\delta)}{(r^k + 1 - \delta)}$.

- Log-linearized aggregate production function:

$$y_t = \Phi\left(\alpha k_t^s + (1-\alpha) l_t + \varepsilon_t^a\right), \tag{4}$$

  where $\Phi = 1 + \frac{\phi}{Y} = 1 + \frac{\text{Steady-state fixed cost}}{Y}$ and α is the capital-share in the production function.[20]

- Effective capital (with one period time-to-build):

$$k_t^s = k_{t-1} + z_t. \tag{5}$$

- Capital utilization:

$$z_t = z_1 r_t^k, \tag{6}$$

where $z_1 = \frac{1-\psi}{\psi}$.

- Capital accumulation equation:

$$k_t = k_1 k_{t-1} + (1 - k_1) i_t + k_2 \varepsilon_t^i, \tag{7}$$

where $k_1 = \frac{1-\delta}{\gamma}$ and $k_2 = \left(1 - \frac{1-\delta}{\gamma}\right)\left(1 + \overline{\beta}\right)\gamma^2 \varphi$.

- Marginal cost:

$$mc_t = (1 - \alpha) w_t + \alpha r_t^k - \varepsilon_t^a. \tag{8}$$

- New-Keynesian Phillips curve (price inflation dynamics):

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 mc_t + \pi_4 \varepsilon_t^p, \tag{9}$$

where $\pi_1 = \frac{\iota_p}{1+\overline{\beta}\iota_p}$, $\pi_2 = \frac{\overline{\beta}}{1+\overline{\beta}\iota_p}$, $\pi_3 = \frac{A}{1+\overline{\beta}\iota_p}\left[\frac{\left(1-\overline{\beta}\xi_p\right)\left(1-\xi_p\right)}{\xi_p}\right]$, and $\pi_4 = \frac{1+\overline{\beta}\iota_p}{1+\overline{\beta}\iota_p}$. The coefficient of the curvature of the Kimball goods market aggregator, included in the definition of A, is fixed in the estimation procedure at $\varepsilon_p = 10$ as in Smets and Wouters (2007).

- Optimal demand for capital by firms:

$$-\left(k_t^s - l_t\right) + w_t = r_t^k. \tag{10}$$

[20]From the zero profit condition in steady-state, it should be noticed that ϕ_p also represents the value of the steady-state price mark-up.

- Wage markup equation:

$$\mu_t^w = w_t - mrs_t = w_t - \left(\sigma_l l_t + \tfrac{1}{1-h/\gamma}\left(c_t - (h/\gamma)\, c_{t-1}\right)\right). \tag{11}$$

- Real wage dynamic equation:

$$w_t = w_1 w_{t-1} + (1-w_1)\left(E_t w_{t+1} + E_t \pi_{t+1}\right) - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu_t^w + \varepsilon_t^w. \tag{12}$$

  where $w_1 = \frac{1}{1+\overline{\beta}}$, $w_2 = \frac{1+\overline{\beta}\iota_w}{1+\overline{\beta}}$, $w_3 = \frac{\iota_w}{1+\overline{\beta}}$, $w_4 = \frac{1}{1+\overline{\beta}}\left[\frac{(1-\overline{\beta}\xi_w)(1-\xi_w)}{\xi_w((\phi_w-1)\varepsilon_w+1)}\right]$ with the curvature of the Kimball labor aggregator fixed at $\varepsilon_w = 10.0$ and a steady-state wage mark-up fixed at $\phi_w = 1.5$ as in Smets and Wouters (2007).

- 1-year term spread:

$$sp_t^{\{4\}} = r_t^{\{4\}} - r_t. \tag{13}$$

- Monetary policy rule:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[r_\pi \pi_t + r_y \hat{y}_t\right] + r_{\Delta y}\Delta\hat{y}_t + r_{sp} sp_t^{\{4\}} + \varepsilon_t^r, \tag{14}$$

  where the output gap is defined as $\hat{y}_t = y_t - \Phi\varepsilon_t^a$ (i.e. the output gap is defined as the deviation of output from its underlying neutral productivity process).

Adaptive learning expectation formation

This part of the appendix provides a brief explanation of how adaptive learning (AL) expectation formation works.[21] A DSGE model can be represented in matrix form as follows:

$$A_0 \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + A_1 \begin{bmatrix} y_t \\ w_t \end{bmatrix} + \sum_{j=1}^{n} A_2^{\{j\}} E_t y_{t+j} + B_0 \epsilon_t = 0,$$

[21] For a detailed explanation see Slobodyan and Wouters (2012a).

where y_t is the vector of endogenous variables at time t and w_t is the exogenous driving force following an AR(1):

$$w_t = \Gamma w_{t-1} + \Pi \epsilon_t,$$

where ϵ_t is the vector of innovations.

Under AL, the expectations of the forward-looking variables, $E_t y_{t+j}$, are defined as linear functions of variables entering in the information set of agents, whose time-varying (learning) coefficients are updated as explained below. Once the expectations of the forward-looking variables, $E_t y_{t+j}$, are computed they are plugged into the matrix representation of the DSGE model to obtain a backward-looking representation of the model as follows

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

where the time-varying matrices μ_t, T_t and R_t are nonlinear functions of structural parameters (entering in matrices A_0, A_1,$A_2^{\{j\}}$ and B_0) together with learning coefficients discussed next.

The perceived law of motion process is generally defined as follows:

$$y_{t+j} = X_t \beta_{t-1}^{\{j\}} + u_{t+j},$$

where y is the vector containing the k forward-looking variables of the model, X is the matrix of the kxr regressors, $\beta^{\{j\}}$ is the vector of the r updating learning coefficients, which includes an intercept, and u is a vector of errors. These errors are linear combinations of the true model innovations, ϵ_t. So, the variance-covariance matrix, $\Sigma = E[u_{t+j} u_{t+j}^T]$, is non-diagonal.

Agents are further assumed to behave as econometricians under AL. In particular, it is assumed that they use a linear projection scheme in which the parameters are updated to form their expectations for each forward-looking variable:

$$E_t y_{t+j} = X_t \beta_{t-1}^{\{j\}}.$$

Thus, agents update their learning coefficient estimates using data up to time $t-1$, but X_t contains contemporaneous values of the regressors. The updating parameter vector, β, which results from stacking all the vectors $\beta^{\{j\}}$, is further assumed to follow an autoregressive process where agents' beliefs are updated through a Kalman filter. This updating process can be represented as in Slobodyan and Wouters (2012a) by the following equation:

$$\beta_t - \bar{\beta} = F(\beta_{t-1} - \bar{\beta}) + v_t,$$

where F is a diagonal matrix with the learning parameter $\mid \rho \mid \leq 1$ on the main diagonal and v_t are i.i.d. errors with variance-covariance matrix V. Notice that the learning parameter ρ plays a role very similar to the *constant gain* parameter used in the recursive least-squares learning.

The Kalman filter updating and transition equations for the belief coefficients and the corresponding covariance matrix are given by

$$\beta_{t|t} = \beta_{t|t-1} + R_{t|t-1} X_{t-1} \Big[\Sigma + X_{t-1}^T R_{t|t-1}^{-1} X_{t-1} \Big]^{-1} \Big(y_t - X_{t-1} \beta_{t|t-1} \Big),$$

with $(\beta_{t+1|t} - \bar{\beta}) = F(\beta_{t|t} - \bar{\beta})$. $\beta_{t|t-1}$ is the estimate of β using the information up to time $t-1$, and $R_{t|t-1}$ is the mean squared error associated with $\beta_{t|t-1}$. Therefore, the updated learning vector $\beta_{t|t}$ is equal to the previous one, $\beta_{t|t-1}$, plus a correction term that depends on the forecast error, $\Big(y_t - X_{t-1} \beta_{t|t-1} \Big)$. Moreover, the mean squared error, $R_{t|t}$, associated with this updated estimate, $\beta_{t|t}$, is given by

$$R_{t|t} = R_{t|t-1} - R_{t|t-1} X_{t-1} \Big[\Sigma + X_{t-1}^T R_{t|t-1}^{-1} X_{t-1} \Big]^{-1} X_{t-1}^T R_{t|t-1}^{-1},$$

with $R_{t+1|t} = F R_{t|t} F^T + V$.

Table A. Model parameter description

φ	Elasticity of the cost of adjusting capital
h	External habit formation
σ_c	Inverse of the elasticity of intertemporal substitution in utility function
σ_l	Inverse of the elasticity of labor supply with respect to the real wage
ξ_p	Calvo probability that measures the degree of price stickiness
ξ_w	Calvo probability that measures the degree of wage stickiness
ι_w	Degree of wage indexation to past wage inflation
ι_p	Degree of price indexation to past price inflation
ψ	Elasticity of capital utilization adjustment cost
Φ	One plus steady-state fixed cost to total cost ratio (price mark-up)
r_π	Inflation coefficient in monetary policy rule
ρ_r	Smoothing coefficient in monetary policy rule
r_Y	Output gap coefficient in monetary policy rule
$r_{\Delta Y}$	Output gap growth coefficient in monetary policy rule
π	Steady-state rate of inflation
$100(\beta^{-1}-1)$	Steady-state rate of discount
l	Steady-state labor
γ	One plus steady-state rate of output growth
$\bar{r}^{\{4\}}$	Mean of the 1-year maturity yield
α	Capital share in production function
ρ	Updating beliefs coefficient

Table A. (*Continued*)

σ_a	Standard deviation of productivity innovation
σ_b	Standard deviation of risk premium innovation
σ_g	Standard deviation of exogenous spending innovation
σ_i	Standard deviation of investment-specific innovation
σ_R	Standard deviation of monetary policy rule innovation
σ_p	Standard deviation of price mark-up innovation
σ_w	Standard deviation of wage mark-up innovation
$\sigma^{\{4\}}$	Standard deviation of one year term premium innovation
ρ_a	Autoregressive coefficient of productivity shock
ρ_b	Autoregressive coefficient of risk premium shock
ρ_g	Autoregressive coefficient of exogenous spending shock
ρ_i	Autoregressive coefficient of investment-specific shock
ρ_R	Autoregressive coefficient of policy rule shock
ρ_p	Autoregressive coefficient of price mark-up shock
ρ_w	Autoregressive coefficient of wage mark-up shock
$\rho^{\{4\}}$	Autoregressive coefficient of one year term premium shock
μ_p	Moving-average coefficient of price mark-up shock
μ_w	Moving-average coefficient of wage mark-up shock
ρ_{ga}	Correlation coefficient between productivity and exogenous spending shocks

Appendix Chapter 2

Supplementary appendix Part II

For a detailed list of the model core variables see the Appendix 1.

Table A.1.A: Priors and estimated posteriors of the structural parameters

	Priors			Posteriors					
				AL model			RE model		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%
log data density					563.99			243.34	
φ: cost of adjusting capital	Normal	4.00	1.50	3.23	2.80	3.69	6.47	6.24	6.74
h: habit formation	Beta	0.70	0.10	0.28	0.24	0.32	0.17	0.14	0.19
σ_l : Frisch elasticity	Normal	2.00	0.75	1.43	1.30	1.62	2.20	2.13	2.27
ξ_p: price Calvo probability	Beta	0.50	0.10	0.56	0.52	0.60	0.936	0.932	0.939
ξ_w: wage Calvo probability	Beta	0.50	0.10	0.77	0.74	0.79	0.81	0.78	0.83
ι_w: wage indexation	Beta	0.50	0.15	0.52	0.43	0.64	0.48	0.43	0.54
ι_p: price indexation	Beta	0.50	0.15	0.37	0.31	0.46	0.12	0.09	0.15
ψ: capital utilization adjusting cost	Beta	0.50	0.15	0.36	0.29	0.44	0.53	0.49	0.58
Φ : steady state price mark-up	Normal	1.25	0.12	1.58	1.49	1.66	1.47	1.42	1.52
r_π: policy rule inflation	Normal	1.50	0.25	1.34	1.28	1.43	1.41	1.37	1.47
ρ_r: policy rule smoothing	Beta	0.75	0.10	0.84	0.82	0.85	0.82	0.80	0.83
r_y: policy rule output gap	Normal	0.12	0.05	0.002	-0.001	0.004	0.009	0.007	0.010
$r_{\Delta y}$: policy rule output gap growth	Normal	0.12	0.05	0.09	0.08	0.11	0.07	0.06	0.08
r_{sp}: policy rule term spread	Normal	0.12	0.05	0.10	0.08	0.13	0.33	0.29	0.36
π: steady-state inflation	Gamma	0.62	0.10	0.67	0.63	0.72	0.70	0.67	0.72
$100(\beta^{-1} - 1)$: steady-state rate of disc.	Gamma	0.25	0.10	0.32	0.26	0.37	0.20	0.17	0.24
l: steady-state labor	Normal	0.00	2.00	4.30	4.03	4.51	5.18	4.86	5.61
γ: one plus st-state rate of output growth	Normal	0.40	0.10	0.56	0.55	0.57	0.30	0.28	0.32
α: capital share	Normal	0.30	0.05	0.17	0.16	0.20	0.27	0.24	0.30
ρ: learning parameter	Beta	0.50	0.28	0.71	0.67	0.75	-	-	-

Table A.1.B: Priors and estimated posteriors of the term structure parameters

	Priors			Posteriors					
				AL			RE		
$\bar{r}^{\{4\}}$: steady-state 1-year rate	Normal	1.50	2.00	1.24	1.15	1.32	1.24	1.22	1.26
$\bar{r}^{\{12\}}$: steady-state 3-year rate	Normal	1.50	2.00	1.52	1.45	1.60	1.26	1.24	1.29
$\bar{r}^{\{20\}}$: steady-state 5-year rate	Normal	1.50	2.00	1.60	1.56	1.69	1.36	1.28	1.41
$\bar{r}^{\{28\}}$: steady-state 7-year rate	Normal	1.50	2.00	1.65	1.61	1.71	1.47	1.45	1.49
$\bar{r}^{\{40\}}$: steady-state 10-year rate	Normal	1.50	2.00	1.74	1.69	1.80	1.56	1.54	1.59

Table A.1.C: Priors and estimated posteriors of the structural shock process parameters

	Priors			Posterior					
				AL model			RE model		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%
σ_a: Std. dev. productivity innovation	Invgamma	0.10	2.00	0.38	0.34	0.41	0.39	0.35	0.41
σ_b: Std. dev. risk premium innovation	Invgamma	0.10	2.00	0.80	0.73	0.87	0.022	0.019	0.025
σ_g: Std. dev. exogenous spending innovation	Invgamma	0.10	2.00	0.39	0.35	0.44	0.43	0.39	0.47
σ_i: Std. dev. investment innovation	Invgamma	0.10	2.00	1.23	1.15	1.30	0.49	0.45	0.53
σ_R: Std. dev. monetary policy innovation	Invgamma	0.10	2.00	0.12	0.11	0.13	0.12	0.11	0.14
σ_p: Std. dev. price mark-up innovation	Invgamma	0.10	2.00	0.18	0.16	0.21	0.19	0.17	0.21
σ_w: Std. dev. wage mark-up innovation	Invgamma	0.10	2.00	0.79	0.72	0.86	0.31	0.27	0.35
$\sigma_{\eta\{2\}}$: Std. dev. 2-quarter yield innovation	Invgamma	0.10	2.00	0.55	0.48	0.61	0.04	0.03	0.06
$\sigma_{\eta\{3\}}$: Std. dev. 3-quarter yield innovation	Invgamma	0.10	2.00	0.05	0.03	0.08	0.05	0.04	0.06
$\sigma_{\eta\{4\}}$: Std. dev. 1-year yield innovation	Invgamma	0.10	2.00	0.33	0.30	0.37	0.33	0.30	0.36
$\sigma_{\eta\{12\}}$: Std. dev. 3-year yield innovation	Invgamma	0.10	2.00	0.34	0.30	0.37	1.20	1.13	1.24
$\sigma_{\eta\{20\}}$: Std. dev. 5-year yield innovation	Invgamma	0.10	2.00	0.04	0.03	0.06	2.02	1.96	2.08
$\sigma_{\eta\{28\}}$: Std. dev. 7-year yield innovation	Invgamma	0.10	2.00	0.56	0.51	0.61	2.79	2.71	2.88
$\sigma_{\eta\{40\}}$: Std. dev. 10-year yield innovation	Invgamma	0.10	2.00	1.74	1.59	1.87	3.20	3.12	3.27

Table A.1.C: (Continued)

	Priors			Posterior					
				AL model			RE model		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%
ρ_a: Autoregressive coef. productivity shock	Beta	0.50	0.20	0.91	0.85	0.95	0.87	0.85	0.89
ρ_b: Autoregressive coef. risk-premium shock	Beta	0.50	0.20	0.87	0.85	0.89	0.945	0.941	0.950
ρ_g: Autoregressive coef. exog. spending shock	Beta	0.50	0.20	0.97	0.95	0.98	0.95	0.94	0.96
ρ_i: Autoregressive coef. investment shock	Beta	0.50	0.20	0.96	0.94	0.98	0.37	0.32	0.42
ρ_R: Autoregressive coef. monetary policy shock	Beta	0.50	0.20	0.96	0.94	0.99	0.22	0.19	0.26
ρ_p: Autoregressive coef. price markup shock	Beta	0.50	0.20	0.90	0.82	0.94	0.93	0.91	0.94
ρ_w: Autoregressive coef. wage markup shock	Beta	0.50	0.20	0.95	0.93	0.97	0.68	0.64	0.74
μ_p: MA coef. price markup shock	Beta	0.50	0.20	0.58	0.49	0.68	0.92	0.91	0.93
μ_w: MA coef. wage markup shock	Beta	0.50	0.20	0.30	0.24	0.35	0.55	0.49	0.62
ρ_{ga}: Interact. betw. product. and spending shocks	Beta	0.50	0.25	0.43	0.33	0.51	0.55	0.48	0.64

Table A.1.C: (Continued)

	Priors			Posterior					
				AL model			RE model		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%
$\rho^{\{2\}}$: Autoregressive coef. 2-quarter yield shock	Beta	0.50	0.20	0.945	0.937	0.954	0.45	0.36	0.52
$\rho^{\{3\}}$: Autoregressive coef. 3-quarter yield shock	Beta	0.50	0.20	0.44	0.36	0.54	0.54	0.47	0.60
$\rho^{\{4\}}$: Autoregressive coef. 4-quarter yield shock	Beta	0.50	0.20	0.89	0.86	0.92	0.75	0.74	0.76
$\rho^{\{12\}}$: Autoregressive coef. 12-quarter yield shock	Beta	0.50	0.20	0.88	0.85	0.92	0.80	0.74	0.84
$\rho^{\{20\}}$: Autoregressive coef. 20-quarter yield shock	Beta	0.50	0.20	0.86	0.82	0.89	0.88	0.84	0.92
$\rho^{\{28\}}$: Autoregressive coef. 28-quarter yield shock	Beta	0.50	0.20	0.84	0.77	0.91	0.73	0.67	0.81
$\rho^{\{40\}}$: Autoregressive coef. 40-quarter yield shock	Beta	0.50	0.20	0.89	0.84	0.96	0.80	0.76	0.83
$\rho^{\{2\}}{}_{\xi}$: Interact. betw. 1- and 2-quarter yield shocks	Beta	0.50	0.25	0.68	0.62	0.74	0.57	0.49	0.68
$\rho^{\{3\}}{}_{\xi}$: Interact. betw. 1- and 3-quarter yield shocks	Beta	0.50	0.25	0.49	0.29	0.72	0.60	0.54	0.65
$\rho^{\{4\}}{}_{\xi}$: Interact. betw. 1- and 4-quarter yield shocks	Beta	0.50	0.25	0.71	0.67	0.76	0.58	0.50	0.68
$\rho^{\{12\}}{}_{\xi}$: Interact. betw. 1- and 12-quarter yield shocks	Beta	0.50	0.25	0.64	0.58	0.69	0.58	0.51	0.65
$\rho^{\{20\}}{}_{\xi}$: Interact. betw. 1- and 20-quarter yield shocks	Beta	0.50	0.25	0.57	0.51	0.66	0.62	0.54	0.71
$\rho^{\{28\}}{}_{\xi}$: Interact. betw. 1- and 28-quarter yield shocks	Beta	0.50	0.25	0.55	0.43	0.65	0.52	0.46	0.57
$\rho^{\{40\}}{}_{\xi}$: Interact. betw. 1- and 40-quarter yield shocks	Beta	0.50	0.25	0.38	0.30	0.48	0.57	0.52	0.63

Table A.1.C: Estimated parameters describing the deviations of model and SPF expectations

	Priors			Posterior					
				AL model			RE model		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%
$\sigma_{\pi}^{\{1\}}$: Std. dev. 1-q-a inflation expect. innov.	Invgamma	0.10	2.00	0.08	0.07	0.09	0.073	0.066	0.081
$\sigma_{\pi}^{\{2\}}$: Std. dev. 2-q-a inflation expect. innov.	Invgamma	0.10	2.00	0.07	0.06	0.08	0.060	0.054	0.066
$\sigma_{\pi}^{\{3\}}$: Std. dev. 3-q-a inflation expect. innov.	Invgamma	0.10	2.00	0.067	0.060	0.074	0.059	0.053	0.065
$\sigma_{\pi}^{\{4\}}$: Std. dev. 4-q-a inflation expect. innov.	Invgamma	0.10	2.00	0.21	0.19	0.23	0.061	0.055	0.067
$\sigma_{\Delta c}^{\{1\}}$: Std. dev. 1-q-a cons. growth expect. innov.	Invgamma	0.10	2.00	0.56	0.50	0.60	0.17	0.16	0.19
$\sigma_{\Delta c}^{\{2\}}$: Std. dev. 2-q-a cons. growth expect. innov.	Invgamma	0.10	2.00	0.13	0.11	0.14	0.15	0.13	0.16
$\sigma_{\Delta c}^{\{3\}}$: Std. dev. 3-q-a cons. growth expect. innov.	Invgamma	0.10	2.00	0.11	0.09	0.12	0.16	0.14	0.17
$\sigma_{\Delta c}^{\{4\}}$: Std. dev. 4-q-a cons. growth expect. innov.	Invgamma	0.10	2.00	0.08	0.07	0.09	0.12	0.11	0.14
$\rho_{\pi}^{\{1\}}$: persist. 1-q-a inflation expect. shock	Beta	0.50	0.20	0.79	0.70	0.86	0.53	0.46	0.60
$\rho_{\pi}^{\{2\}}$: persist. 2-q-a inflation expect. shock	Beta	0.50	0.20	0.92	0.88	0.96	0.39	0.33	0.45
$\rho_{\pi}^{\{3\}}$: persist. 3-q-a inflation expect. shock	Beta	0.50	0.20	0.87	0.82	0.92	0.52	0.46	0.58
$\rho_{\pi}^{\{4\}}$: : persist. 4-q-a inflation expect. shock	Beta	0.50	0.20	0.88	0.83	0.93	0.66	0.61	0.70
$\rho_{\Delta c}^{\{1\}}$: persist. 1-q-a cons. growth expect. shock	Beta	0.50	0.20	0.94	0.90	0.98	0.88	0.85	0.91
$\rho_{\Delta c}^{\{2\}}$: persist. 2-q-a cons. growth expect. shock	Beta	0.50	0.20	0.71	0.64	0.80	0.93	0.90	0.95
$\rho_{\Delta c}^{\{3\}}$: persist. 3-q-a cons. growth expect. shock	Beta	0.50	0.20	0.32	0.24	0.38	0.92	0.88	0.95
$\rho_{\Delta c}^{\{4\}}$: persist. 4-q-a cons. growth expect. shock	Beta	0.50	0.20	0.68	0.60	0.76	0.92	0.90	0.95
μ_{π}: persistence of inflation expt. rule	Beta	0.80	0.15	0.971	0.968	0.974	–	–	–
μ_{c}: persistence of consumption expt. rule	Beta	0.90	0.15	0.97	0.91	1.00	–	–	–

Table A.2. Selected parameter estimates for two alternative cases

	AL model		RE model	
	SlW Taylor rule	RT inflation	SlW Taylor rule	RT inflation
Number of observables	20	21	20	21
h: habit formation	0.20	0.55	0.84	0.84
	(0.14,0.24)	(0.52,0.58)	(0.82,0.86)	(0.83,0.85)
φ: cost of adjusting capital	1.71	1.49	7.27	9.43
	(1.43,2.03)	(1.38,1.61)	(6.54,8.00)	(8.35,10.57)
ψ: capital utilization adjusting cost	0.36	0.87	0.39	0.67
	(0.30,0.41)	(0.82,0.92)	(0.35,0.44)	(0.58,0.76)
ξ_p: price Calvo probability	0.52	0.44	0.94	0.92
	(0.48,0.56)	(0.40,0.48)	(0.93,0.95)	(0.91,0.94)
ξ_w: wage Calvo probability	0.38	0.31	0.77	0.78
	(0.31,0.44)	(0.27,0.35)	(0.71,0.83)	(0.75,0.82)
ι_p: price indexation	0.21	0.30	0.10	0.11
	(0.14,0.28)	(0.22,0.37)	(0.07,0.12)	(0.07,0.15)
ι_w: wage indexation	0.62	0.37	0.45	0.49
	(0.41,0.81)	(0.31,0.45)	(0.32,0.54)	(0.26,0.68)
ρ_p: persistence of price markup shock	0.93	0.95	0.86	0.91
	(0.90,0.96)	(0.93,0.98)	(0.82,0.89)	(0.88,0.95)
ρ_w: persistence of wage markup shock	0.96	0.95	0.79	0.73
	(0.93,0.98)	(0.92,0.98)	(0.68,0.90)	(0.66,0.81)

Table A.2. (*Continued*)

	AL model		RE model	
	SlW Taylor rule	RT inflation	SlW Taylor rule	RT inflation
Number of observables	20	21	20	21
μ_p: MA coef. price markup shock	0.57	0.50	0.83	0.90
	(0.51,0.64)	(0.43,0.56)	(0.79,0.87)	(0.86,0.94)
μ_w: MA coef. wage markup shock	0.31	0.31	0.68	0.63
	(0.19,0.44)	(0.27,0.36)	(0.52,0.82)	(0.54,0.73)
μ_π: persist. of inflat. rule expect.	0.982	0.999		
	(0.980,0.987)	—		
μ_c: persist. of conspt. rule expect.	0.97	0.999		
	(0.91,1.00)	—		
log data density	545.01	566.27	220.24	222.86

Notes: parameter notation and standard deviation in parentheses

Appendix Chapter 4

Supplementary appendix Part IV

The log-linearised system of equations

We use equations (1), (4), (5), (7), (11), (12), (15) and (16) log-linearise it to form a system of 8 equations and 8 variables (y_t, w_t, r_t, tr_t, i_t, k_t, c_t, n_t) in addition to the four shock processes.

We start from the production function in equation (1), in detrended terms $y_t = \exp(z_t) n_t^{1-\alpha} k_{t-1}^{\alpha}$, and take logs:

$$\ln(y_t) = ln\left(z_t\right) + (1-\alpha)\ln(n) + \alpha \ln(k_{t-1})$$

The taylor expansion around the steady state is:

$$\ln(y^{ss}) + \left(\frac{y_t - y^{ss}}{y^{ss}}\right) = \ln\left(\exp(z_t)\right) + \left(\frac{z_t - z^{ss}}{z^{ss}}\right)$$
$$+ (1-\alpha)\ln(n^{ss}) + (1-\alpha)\left(\frac{n_t - n^{ss}}{n^{ss}}\right) + \alpha\ln(k^{ss}) + \alpha\left(\frac{k_{t-1} - k^{ss}}{k^{ss}}\right),$$

where the term on the right hand side $\ln(y^{ss})$ is equal to $\ln(\exp(z_t)) + (1-\alpha)\ln(n_t) + \alpha\ln(k_{t-1})$, therefore we have:

$$\left(\frac{y_t - y^{ss}}{y^{ss}}\right) = \left(\frac{z_t - z^{ss}}{z^{ss}}\right) + (1-\alpha)\left(\frac{n_t - n^{ss}}{n^{ss}}\right) + \alpha\left(\frac{k_{t-1} - k^{ss}}{k^{ss}}\right)$$

We can define $\hat{y}_t = \left(\dfrac{y_t - y^{ss}}{y^{ss}} \right)$, and our "tilde" equation becomes :

$$\hat{y}_t = z_t + (1-\alpha)\hat{n}_t + \alpha \hat{k}_{t-1} \tag{1}$$

The linearised demand for labor in equation (4):

Taking logs:

$$\ln(w_t) = \ln\left(\exp(z_t)\right) + \ln(1-\alpha) - \alpha \ln(n_t) + \alpha \ln(k_{t-1}),$$

where the taylor expansion around the steady state is:

$$\ln(w^{ss}) + \left(\frac{w_t - w^{ss}}{w^{ss}} \right) = \ln\left(\exp(z_t)\right) + \left(\frac{z_t - z^{ss}}{z^{ss}} \right)$$
$$+ \ln(1-\alpha) - \alpha \ln(n^{ss}) - \alpha \left(\frac{n_t - n^{ss}}{n^{ss}} \right) + \alpha \ln(k^{ss}) + \alpha \left(\frac{k_{t-1} - k^{ss}}{k^{ss}} \right),$$

where the term on the right hand side $\ln(w^{ss})$ is equal to $\ln\left(\exp(z_t)\right) + \ln(1-\alpha) - \alpha \ln(n_t) + \alpha \ln(k_{t-1})$, we have that

$$\left(\frac{w_t - w^{ss}}{w^{ss}} \right) = \left(\frac{z_t - z^{ss}}{z^{ss}} \right) - \alpha \left(\frac{n_t - n^{ss}}{n^{ss}} \right) + \alpha \left(\frac{k_{t-1} - k^{ss}}{k^{ss}} \right),$$

and under "tilde" becomes:

$$\hat{w}_t = z_t + \alpha(\hat{k}_{t-1} - \hat{n}_t) \tag{2}$$

The linearised demand for capital (5) is similar to the previous expresion, thus:

$$\hat{r}_t = z_t + (1-\alpha)(\hat{n}_t - \hat{k}_{t-1}) \tag{3}$$

The linearised equation of the Goverment's budget (7), after taking logs:

$$\ln(g_t) = \ln\left(w_t \tau_t^w n_t + r_t \tau_t^r k_{t-1} - tr_t\right),$$

where the taylor expansion around the steady state is:

$$\ln(g^{ss}) + \left(\frac{g_t - g^{ss}}{g^{ss}}\right) = \ln\left(w_t \tau_t^w n_t + r_t \tau_t^r k_{t-1} - tr_t\right) + \frac{\gamma^{nss} n^{ss}\left(w_t - w^{ss}\right)}{w^{ss}\tau^{wss} n^{ss} + r^{ss}\tau^{rss} k^{ss} - tr^{ss}}$$

$$+ \frac{w^{ss} n^{ss}\left(\tau_t^w - \tau^{nss}\right) + w^{ss}\gamma^{nss}\left(w_t - w^{ss}\right) + \tau^{rss} k^{ss}\left(r_t - r^{ss}\right)}{w^{ss}\tau^{wss} n^{ss} + r^{ss}\tau^{rss} k^{ss} - tr^{ss}}$$

$$+ \frac{r^{ss} k^{ss}\left(\tau_t^r - \tau^{rss}\right) + r^{ss}\gamma^{rss}\left(k_{t-1} - k^{ss}\right) + \left(tr_t - tr^{ss}\right)}{w^{ss}\tau^{wss} n^{ss} + r^{ss}\tau^{rss} k^{ss} - tr^{ss}},$$

where we know that the expresion in the denominator is equal to g^{ss}, multiplying and dividing each term by its value in steady state:

$$g_{ss}\hat{g}_t = w_{ss}\tau_{ss}^w n_{ss}(\hat{w}_t + \hat{\tau}_t^w + \hat{n}_t) + r_{ss}\tau_{ss}^r k_{ss}(\hat{r}_t + \hat{\tau}_t^r + \hat{k}_{t-1}) - \frac{tr^{ss}\hat{tr}_t}{g_{ss}} \tag{4}$$

The linearised household budget constraint (11) can be rewritten using the production function and the goverment budget constraint, as $y_t = c_t^p + i_t + g_t$, taking logs and operating, we reach:

$$y_{ss}\hat{y}_t = c_{ss}\hat{c}_t^p + i_{ss}\hat{\imath}_t + g_{ss}\hat{g}_t \tag{5}$$

The linearised capital law of motion (12):

$$\hat{i}_t = \frac{k^{ss}}{i^{ss}}\gamma\hat{k}_{t-1} - \frac{k^{ss}}{i^{ss}}\left(1-\delta\right)\hat{k}_{t-1} \tag{6}$$

The linearised Euler equation (15): For simplicitiy, recall $\Omega = \beta/(1+\gamma)$ and $R_{t+1} =$

$(1-\tau^r_{t+1})\alpha\exp(z_t)k_t^{\alpha-1}n_{t+1}^{1-\alpha}+1-\delta$, and taking logs:

$$\ln\left(\frac{\Omega}{c_t^p+\eta g_t}\right)=\ln(\frac{\beta}{c_t^p+\eta g_t})+\ln(R_{t+1}),$$

then we have $\ln\Omega-\ln(c_t^p+\eta g_t)=\ln\beta-\ln(c_{t+1}^p+\eta g_{t+1})+\ln(R_{t+1})$, which becomes:

$$\begin{aligned}-\ln(c_t^p+\eta g_t)-\left(\frac{c_t^p+c^{pss}}{c_{ss}^p+g_{ss}}\right)-\eta\left(\frac{g_t+g^{ss}}{c_{ss}^p+g_{ss}}\right)&=-\ln(c_{t+1}^p+\eta g_{t+1})\\&-\left(\frac{c_{t+1}^p+c^{pss}}{c_{ss}^p+g_{ss}}\right)-\eta\left(\frac{g_{t+1}+g^{ss}}{c_{ss}^p+g_{ss}}\right)+\frac{R_{t+1}+R^{ss}}{R_{ss}}+\ln(R_{t+1})\end{aligned}$$

Multipying and dividing by c_{ss}^p and g_{ss}, we have:

$$-\frac{c_{ss}c_t^p+g_{ss}\eta g_t}{c_{ss}^p+g_{ss}}=-\frac{c_{ss}c_{t\,t+1}^p+g_{ss}\eta g_{t+1}}{c_{ss}^p+g_{ss}}+R_{t+1},$$

where $R_{t+1}=\left(\frac{r_{ss}}{r_{ss}+1-\delta}\right)(z_{t+1}+(1-\alpha)(n_{t+1}-k_t)-\tau^r_{ss}\tau^r_t)$, and finally our linearised equation becomes:

$$\frac{c_{ss}\hat{c}_t^p+g_{ss}\eta\hat{g}_t}{c_{ss}^p+g_{ss}}=\frac{c_{ss}\hat{c}_{t\,t+1}^p+g_{ss}\eta\hat{g}_{t+1}}{c_{ss}^p+g_{ss}}-\left(\frac{r_{ss}}{r_{ss}+1-\delta}\right)\left(\hat{z}_{t+1}+(1-\alpha)(\hat{n}_{t+1}-\hat{k}_t)-\tau^r_{ss}\hat{\tau}^r_{t+1}\right).\tag{7}$$

The linearised labour market equilibrium equation (16). Taking logs we have:

$$\ln\left(\frac{\varphi}{1-n_t}\right)=\ln(w_t)+\ln(1-\tau_t^n)-\ln\left(c_t+\eta g_t\right)$$

The taylor expansion around the steady state is:

$$\ln(n^{ss}) + \varphi\left(\frac{n_t - n^{ss}}{1-n^{ss}}\right) = \ln(w_t) + \left(\frac{w_t - w^{ss}}{w^{ss}}\right) + \ln(1-\tau_t^n) - \left(\frac{\tau_t^n - \tau^{nss}}{\tau^{nss}}\right)$$
$$- \ln\left(c_t^p + \eta g_t\right) - \frac{\eta g^{ss}}{c^{pss}+g^{ss}}\left(\frac{c_t^p - c^{ss}}{c^{ss}}\right) - \frac{c^{pss}\eta}{c^{pss}+g^{ss}}\left(\frac{g_t - g^{ss}}{g^{ss}}\right),$$

we multiply the left hand side by n^{ss} to obtain $\widetilde{n}$, replace $\widetilde{w}_t$, and reach:

$$\hat{n}_t = \left(\frac{1-n^{ss}}{n^{ss}+\alpha(1-n^{ss})}\right)\left(\hat{z}_t - \alpha\hat{k}_{t-1} - \left(\frac{c^{pss}\hat{c}_t^p + g^{ss}\eta\hat{g}_t}{c^{pss}+g^{ss}}\right) - \tau^{wss}\hat{\tau}_t^w\right) \tag{8}$$

The reduced system of equations

We now show the steps for obtaining the reduced form system of the model, this is, as a function of the state variable and the shocks, it is important to follow a strategy in order to reduce the time devoted to algebra. The to obtain an expresion of consumption as a function of the state variable and the shocks, for doing so we use the resource constraint of the economy (5-A, corresponding to appendix) combined with equations (1) and (6).

$$y_{ss}\left(\hat{z}_t + (1-\alpha)\hat{n}_t + \alpha\hat{k}_{t-1}\right) = c_{ss}\hat{c}_t^p + \gamma k^{ss}\hat{k}_t - k^{ss}(1-\delta)\hat{k}_{t-1} + g_{ss}\hat{g}_t.$$

Next, we replace $\hat{n}_t$ using equation (8-A), and we obtain:

$$c_{ss}\hat{c}_t^p\left(1 + \frac{n_1}{c^{pss}+g^{ss}}\right) = (y_{ss}+n_1)\hat{z}_t + (n_1\alpha + y_{ss}\alpha + k_{ss}(1-\delta)k^{ss})\hat{k}_{t-1}$$
$$- \gamma k^{ss}\hat{k}_t - g_{ss}\hat{g}_t\left(1 + \frac{\eta n_1}{c^{pss}+g^{ss}}\right) - n_1\tau^{wss}\hat{\tau}_t^w,$$

where $n_1 = (1-\alpha)\left(\dfrac{1-n_{ss}}{n_{ss}+\alpha(1-n_{ss}}\right)$ This equation is going to be crucial in obtaining the reduced form of the model, for the sake of simplicity we rewrite it as:

$$\hat{c}_t^p = z_0\hat{z}_t + k_0\hat{k}_{t-1} - k_1\hat{k}_t - g_0\hat{g}_t - \tau_0^w\hat{\tau}_t^w, \tag{9}$$

where $z_0 = \dfrac{y_{ss}+n_1}{c^{pss}\left(1+\dfrac{n_1}{c^{pss}+g^{ss}}\right)}$, $k_0 = \dfrac{n_1\alpha + y_{ss}\alpha + k_{ss}(1-\delta)k^{ss}}{c^{pss}\left(1+\frac{n_1}{c^{pss}+g^{ss}}\right)}$,

$$k_1 = \frac{\gamma k^{ss}}{c^{pss}\left(1+\dfrac{n_1}{c^{pss}+g^{ss}}\right)},\ g_0 = \frac{1+\dfrac{\eta n_1}{c^{pss}+g^{ss}}}{c^{pss}\left(1+\dfrac{n_1}{c^{pss}+g^{ss}}\right)} \text{ and } \tau_0^w = \frac{n_1\tau^{wss}}{c^{pss}\left(1+\dfrac{n_1}{c^{pss}+g^{ss}}\right)}.$$

Now we operate in the Euler equation and the labor market equilibrium equation to express write the expectations of consumption and labor as a function of the state variable and the shocks. After some algebra, and substituing $\hat{n}_{t+1}$from equation (8-A) in $t+1$ and the shocks in $t+1$ as well, by its autoregresive process (for instance, using$\hat{z}_{t+1} = \rho_z\hat{z}_t$), we can write equation (7-A) as:

$$c_0\hat{c}_t^p = (1+r_1)c_0\hat{c}_{t+1}^p + (\rho_g - 1 + r_1\rho_g)g_1\hat{g}_t - (r_0+r_1)\rho_z\hat{z}_t + (r_0(1-\alpha)-r_1)\hat{k}_t + r_1\rho_w\tau^{wss}\hat{\tau}_t^w + r_0\rho_r\tau^{rss}\hat{\tau}_t^r,$$

where $c_0 = \frac{c^{pss}}{c^{pss}+g_{ss}}$, $g_1 = \frac{\eta g^{ss}}{c^{pss}+g_{ss}}$, $r_0 = \frac{r^{ss}}{r^{ss}+1-\delta}$, and $r_1 = r_0(1-\alpha)n_0$.

Next, we use equation (9-A) in t and $t+1$to substitute $\hat{c}_t$ $\hat{c}_{t+1}$in the previous equation, reaching an equation where only the state variable and the shocks appear:

$$\hat{k}_t = a_1 E_t \hat{k}_{t+1} + a_2 \hat{k}_{t-1} + b_1 \hat{z}_t + b_2 \hat{g}_t + b_3 \hat{\tau}_t^r + b_4 \hat{\tau}_t^w, \tag{10}$$

where $a_0 = -c_0 k_1 - c_1 k_0 - r_0(1-\alpha) + r_1 \alpha$, $c_1 = (1+r_1)c_0$,

$a_1 = -\dfrac{c_1 k_1}{a_0}$, $a_2 = -\dfrac{c_0 k_0}{a_0}$, $b_1 = -\dfrac{c_0 z_0 + c_1 \rho_z z_0 - (r_0 + r_1)\rho_z}{a_0}$,

$b_2 = c_0 g_0 - \rho_g c_1 g_0 + (\rho_g - 1 + r_1 \rho_g) g_1 / a_0$,

$b_3 = \dfrac{r_0 \tau^{rss} \rho_r}{a_0}$, and $b_4 = \dfrac{c_0 \tau_0^w - c_1 \tau_0^w \rho_w + r_1 \tau^{wss} \rho_w}{a_0}$.

The reduced form system of the model is formed with equation (35), (34), (33), (26), (27), (28), (29), (30) and the shock processes (3), (8), (9) and (10). Once the expectation of capital is solved and the shocks are realised, the rest of variables can be obtained.

The k/n ratio and the price of the factors of production

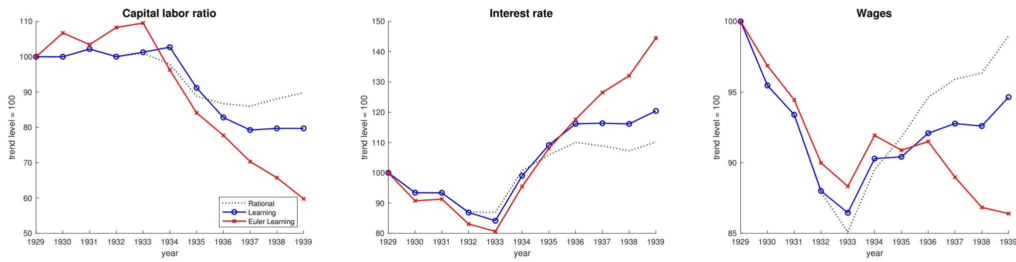


Figure 4: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with circles), model with learning about FOCs (red line with crosses).

Robustness

In this section we shall evaluate the robustness of our results under different specifications of the updating algorithm.

In Figure 5, we compare the results from our benchmark model (constant gain with $\zeta = 0.03$) with the results from simulations with constant gain learning, under different values of the gain parameter. Since in the literature traditional values range from 0.02 to 0.06, we chose those extremes for our sensitivity analysis. The Figure also shows the results from simulations with recursive least squares. We may conclude from this comparison that our results are robust to the specifications of the updating algorithm that are most commonly adopted in the literature.

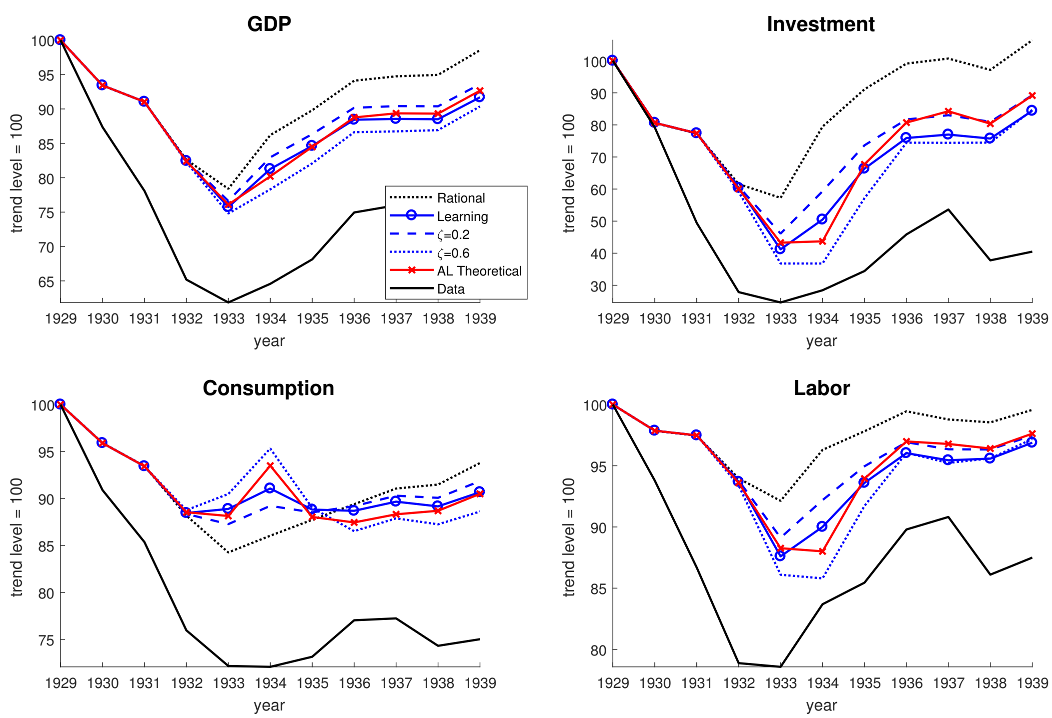


Figure 5: Simulations under different gains. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with cercles) and different gains (dashed lines) and recursive least squares (red line with crosses).

In Figure 6, we check instead how the choice of initial conditions for the learning algory-

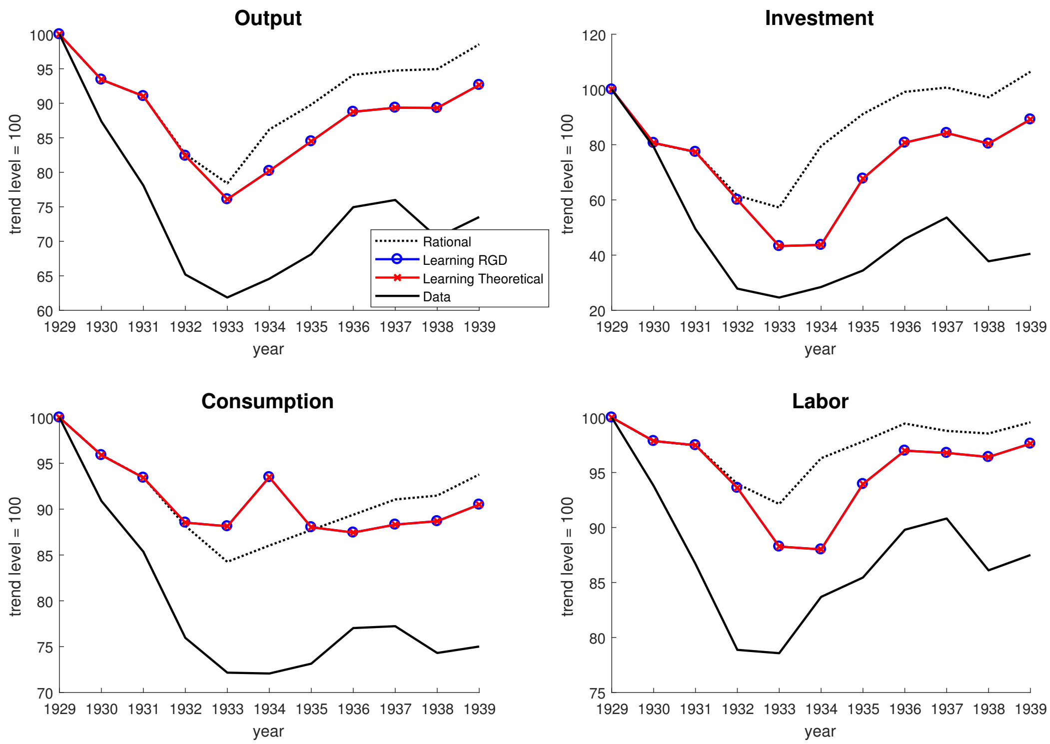


Figure 6: Simulations of the learning model under different intial conditions mechanisms. Model with rational expectations (black-dotted line), learning model initialized with Randomly Generated Data (blue line with cercles) and learning model initialized with the model implied theoretical momments (red line with crosses .

thm, i.e. Equations (24) and (25), impacts on the results. We do so by comparing results from our benchmark model with the results from a model in which the initial value of the variance and covariance matrix, **R**, is the theoretical value implied by the rational expectations model.[22]. The graph shows that the two models almost overlap, meaning that if the simulation for the initial conditions under randomly generated data is sufficiently long, the initial matrix converges to the theoretically-implied one.

Figure 7 shows the results under a 50% higher and lower volatility of the shocks used in the RGD process to verify the sensibility of the results to the variance-covariance intial matrix.

Finally figure 8 shows the model simulation starting from 1920, this is, allowing agents

[22]For the specific representation of this matrix, see [**?**]

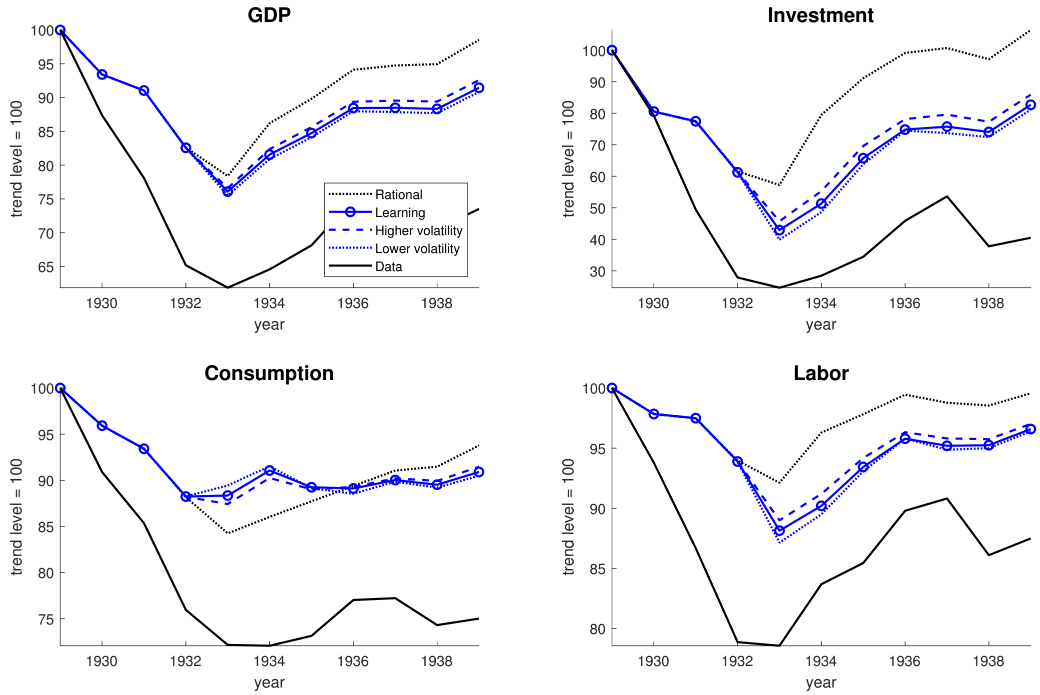


Figure 7: Simulations under different higher (lower) volatility in the RGD process. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with cercles) and higher (blue dashed line) and lower (blue-dotted line) volatility.

to start their learning process 10 years earlier.

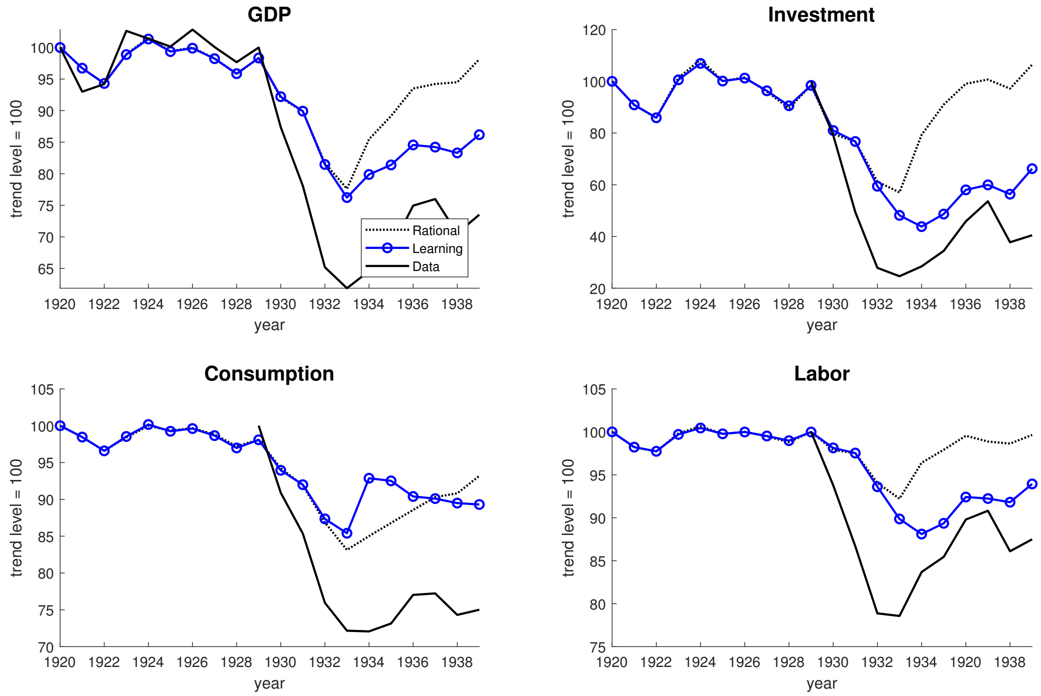


Figure 8: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with markers), data (black line).

Convergence

In this Section, we check the convergence properties of our model. To do so, we run a 1000 periods simulation with large volatility, to impose long periods of stress. We compare the series of output from the model with rational expectations, and that with adaptive learning. Figure 9 shows that the model with adaptive learning and constant gain converges quite fastly to the rational expectation solution and shows similar volatility, whereas the model with FOC-learning à la Milani (Euler learning) tends to amplify the volatility, especially in reaction to large shocks.

Additionally we show the comparison between the RE baseline and the non-linear RE version.

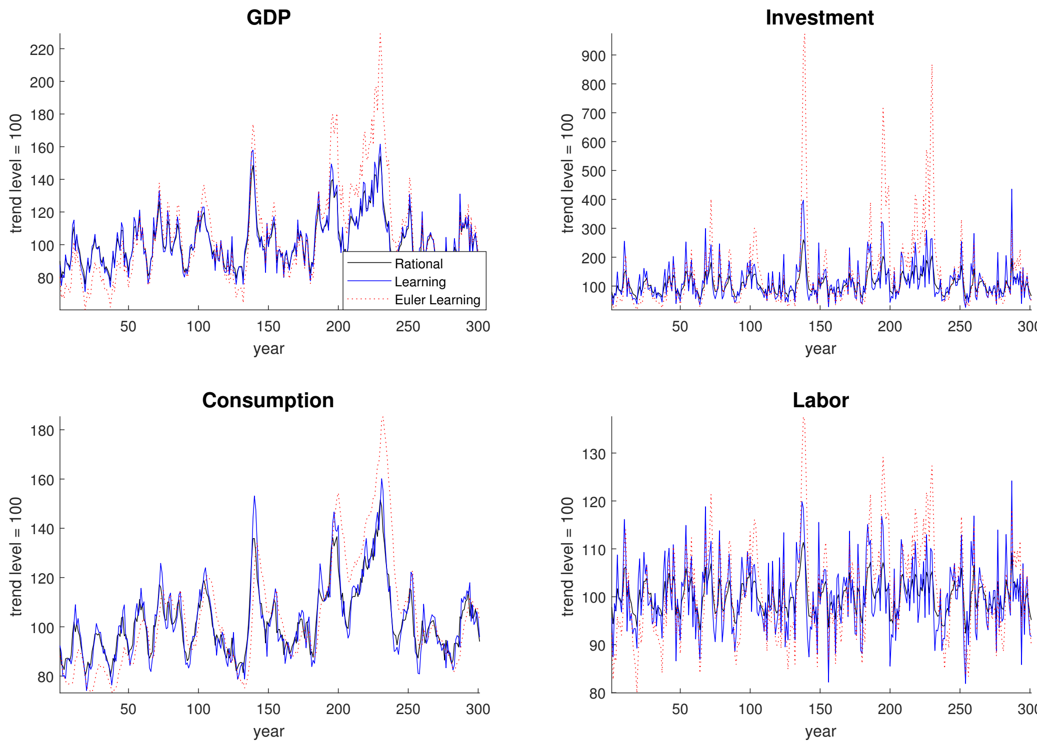


Figure 9: Model dynamics under different expectations, measured in deviations from the steady-state. Model with rational expectations (black line), model with adaptive learning (blue line) and Euler learning (red-dotted line).

Alternative ways of modelling learning

Hybrid expectations

In this Section, we explore an even smaller deviation from rational expectations, i.e. we assume that only a fraction of agents do not know the true values of the actual law of motion, while all the others have rational expectations. This is often referred to in the literature as 'hybrid expectations', for in this context a fraction of agents have forward-looking, rational expectations while the other fraction have backward-looking, adaptive-learning expectations. The hybrid expectations formulation has gained some popularity in recent years (see see Bernanke et al. (2019) and Gertler (2017) for instance) as an easy way to introduce backward-looking agents in general equilibrium models, which brings the persistence of fluctuations in

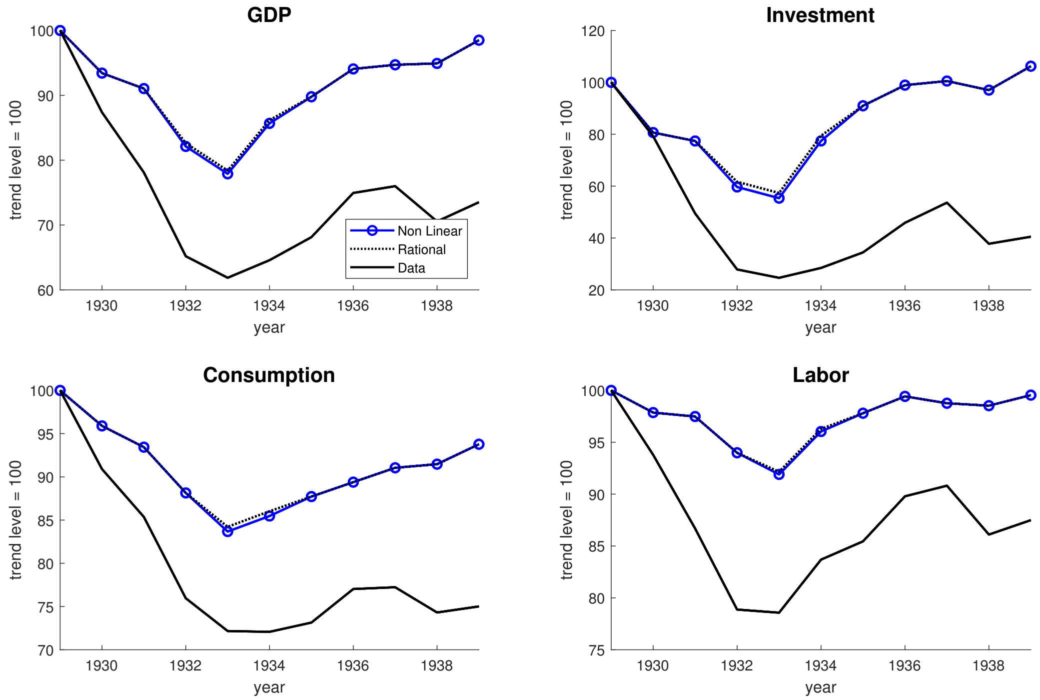


Figure 10: Simulations under different approximations. Model with rational expectations (black-dotted line), model with non-linear rational expectations (blue line with cercles).

the model closer to that in the data. With hybrid expectations, the PLM becomes:

$$E_t\hat{k}_{t+1} = \lambda E_t^{RE}\hat{k}_{t+1} + (1-\lambda)E_t^{AL}\hat{k}_{t+1},$$

where λ represent the fraction of agents whoc know the true values of the ALM. Under this formulation, if λ=1, we have the rational expectation solution, if λ=0, the adaptive learning one. Figure 11 shows the difference in the dynamics of output and consumption for our excercise in Chapter 4, section 4 as the share of rational agents shrinks.

Learning with a constant

In this Section, we extend the benchmark adaptive-learning model by including an intercept in the PLM, to capture the potential bias in the estimation of the coefficients. This extension draws on Slobodyan and Wouters (2012a,b), who introduce it to relax the restriction of agents

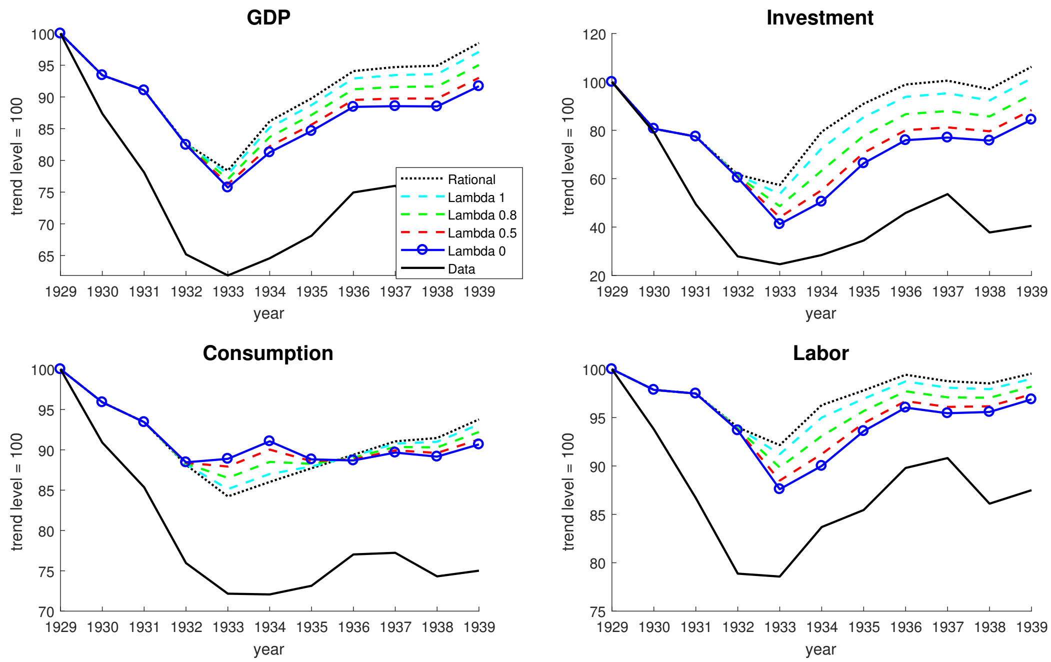


Figure 11: Simulations under different degrees of hybrid expectations. From 1, black dotted line(purely rational) to 0, blue line with cercles (learning).

having a common, constant long-run trend of consumption and inflation. In their paper, the additional term included in the PLM allows expectations to track long-run movements observed in the data, such as the great moderation in inflation during the 1980's. In the case of our model, this term aims more generally at capturing unspecified persistent deviations in the expectations of capital. Adding an intercept to Equation (22) results in the following expression:

$$E_t\hat{k}_{t+1} = \bar{\phi}_{t-1} + \phi_{k,t-1}\hat{k}_t + \phi_{z,t-1}\hat{z}_t + \phi_{g,t-1}\hat{g}_t + \phi_{r,t-1}\hat{\tau}_t^r + \phi_{w,t-1}\hat{\tau}_t^w.$$

By the same token, the learning algorithm becomes:

$$\mathbf{\Phi_t} = \mathbf{\Phi_{t-1}} + \zeta_t \mathbf{R_{t-1}^{-1}} \mathbf{x_{t-1}} (\hat{k}_t - \mathbf{x'_{t-1}} \mathbf{\Phi_{t-1}}), \tag{11}$$

$$\mathbf{R_t} = \mathbf{R_{t-1}} + \zeta_t \big(\mathbf{x_{t-1}} \mathbf{x'_{t-1}} - \mathbf{R_{t-1}}\big), \tag{12}$$

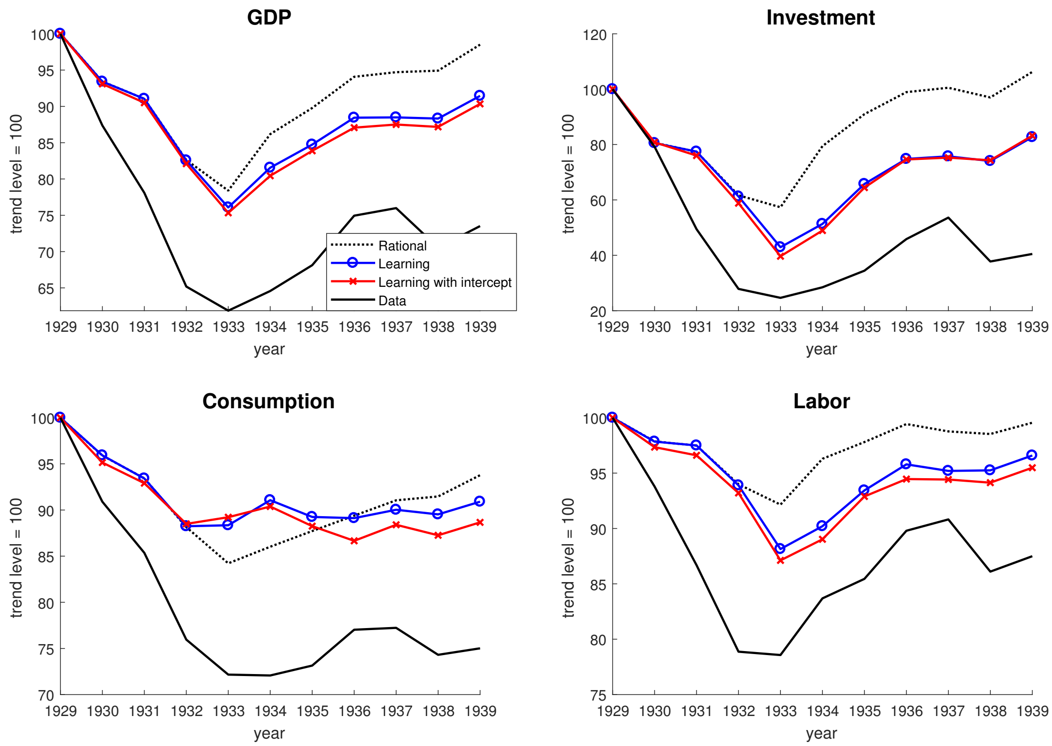


Figure 12: Simulation. Model with rational expectations (black-dotted line), baseline learning model (blue line with cercles) and learning model with an intercept (red line with crosses).

where $\mathbf{x_t}$ is the (column) vector of the state variables $(1, \hat{k}_t, \hat{z}_t, \hat{g}_t, \hat{\tau}_t^r, \hat{\tau}_t^w)$, $\mathbf{R_t}$ is the matrix (of order 6) of variances and covariances of the state variables and $\mathbf{\Phi_t}$ is the (column) vector of the expectation coefficients $(\bar{\phi}, \phi_k, \phi_z, \phi_g, \phi_r, \phi_w)$.

Figure 12 shows the results from simulating the model under the new PLM. The inclusion of an intercept results into a slighly larger drop in output with respect to the baseline learning model.