LEARNING THE HARD WAY: EXPECTATIONS AND THE U.S. GREAT DEPRESSION

Pablo Aguilar and Luca Pensieroso

LIDAM Discussion Paper IRES
2022 / 04

UCLouvain

LIDAM
IRES | Louvain Institute of Data Analysis and Modeling in economics and statistics

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

Pablo Aguilar† Luca Pensieroso‡

February 14, 2022

Abstract

We introduce adaptive learning – a parsimonious, convenient way to model uncertainty – in a dynamic general equilibrium model of the U.S. Great Depression. We show that even the smallest departure from rational expectations increases significantly the data mimicking ability of the model, in particular for what concerns the lack of recovery in detrended GDP after 1933. We conclude that in the case of big, traumatic events like the Great Depression, uncertainty is particularly unfavourable to the recovery phase.

Keywords: Learning, Great Depression, Dynamic general equilibrium, Bounded rationality
JEL Classification: E13, E32, N10

1 Introduction

This research introduces a light form of bounded rationality in a dynamic general equilibrium (DGE) model to explore the role of uncertainty in accounting for the Great Depression of the 1930s in the United States.

*Paper presented at the CEPR Economic History Symposium in Tarragona. We thanks participants for interesting remarks. Giulio Nicoletti made useful remarks on an earlier draft. We thank Jonathan Collard for the excellent research assistance on the very early stage of this project. The usual disclaimers apply.

†Bank of Spain; IRES/LIDAM, Université catholique de Louvain. Email: pablo.aguilar@bde.es

‡IRES/LIDAM, Université catholique de Louvain. Email: luca.pensieroso@uclouvain.be

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 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

[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).

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

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 is particularly unfavourable to the recovery phase, a trait by and large overlooked by the literature so far.

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 complexify 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.

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 favourable 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 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 Prescott (1999), while at the same time stressing the role of uncertainty.

[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).

The literature on adaptive learning was pionereed 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.

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 the PLM has the functional form of the rational expectations law of motion – i.e. the functional form of the minimal state variable solution to the model. Since we view this as the minimum departure from rational expectations, this is our benchmark model, in the spirit of Evans and Honkapohja (2001).[5] In our model, hence, agents know the general form of the minimal state variable solution to the model, but are unable to compute the coefficients associated to it via the method of undetermined coefficient. 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 ot a distribution centred around it, depending on the updating algorithm. An alternative way would be to assume that agents ignore the minimal state variable solution to the model, and formulate *ad hoc* forecasting rules based on information restricted to certain endogenous variables only. This is also known as the 'Euler' approach – as it typically uses variables appearing in the Euler equation. It is the modellisation of choice when introducing adaptive learning in medium-scale New-Keynesian model, like for instance in Aguilar and Vázquez (2019, 2021), Milani (2011), or Slobodyan and Wouters (2012a,b). It represents a larger deviation from the rational expectations hypothesis, since it introduces an additional degree of freedom in the choice of the variables to be included in the formation of expectations. Accordingly, in this type of

[5]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.

models the PLM is no longer restricted to be consistent with the rational expectations decision rule. In an extension to our benchmark, we build a model of the U.S. Great Depression with a Euler-type 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 organised as follows. In Section 2, we present the analytical model, discuss its rational expectations solution and propose our adaptive learning algorithm. 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:

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

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:

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

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 maximises its profits, which delivers the demand for labour and capital

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

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

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 maximise

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

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,

$$\frac{1}{\tilde{c}_t} = \frac{\beta}{1+\gamma} E_t \left(\frac{1}{\tilde{c}_{t+1}} \left(\left(1 - \tau^r_{t+1}\right) (\alpha \exp(z_{t+1}) \tilde{k}_t^{\alpha-1} n_{t+1}^{1-\alpha}) + 1 - \delta \right) \right), \qquad (15)$$

and the equation determining the equilibrium on the labour market,

$$\frac{\varphi}{1-n_t} = (1-\tau^w_t) \frac{(1-\alpha)\exp(z_t)\tilde{k}^{\alpha}_{t-1} n_t^{1-\alpha}}{\tilde{c}_t}, \qquad (16)$$

where $\tilde{c}_t$ ($\tilde{k}_t$) stands for detrended c_t (k_t), that is c_t/x_t (k_t/x_t).

2.2 Expectations

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

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

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

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

$$\hat{\tau}^r_t = \rho_r \hat{\tau}^r_{t-1} + u_t, \qquad (17d)$$

$$\hat{\tau}^w_t = \rho_w \hat{\tau}^w_{t-1} + s_t, \qquad (17e)$$

$$\hat{q}_t = \psi^q_1 \hat{k}_t + \psi^q_2 \hat{k}_{t-1} + \psi^q_3 \hat{z}_t + \psi^q_4 \hat{g}_t + \psi^q_5 \hat{\tau}^r_t + \psi^q_6 \hat{\tau}^w_t, \qquad (17f)$$

with $\hat{q} = \hat{y}, \hat{c}, \hat{n}, \hat{tr}$, and $\hat{y}$ meaning y in log-deviations from steady state. 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 materialises.

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:

$$\hat{k}_t = \chi_k \hat{k}_{t-1} + \chi_z \hat{z}_t + \chi_g \hat{g}_t + \chi_r \hat{\tau}^r_t + \chi_w \hat{\tau}^w_t, \qquad (18)$$

where χ_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 becomes:

$$\hat{k}_t = \phi_k \hat{k}_{t-1} + \phi_z \hat{z}_{t-1} + \phi_g \hat{g}_{t-1} + \phi_r \hat{\tau}^r_{t-1} + \phi_w \hat{\tau}^w_{t-1} + \mu_t + \varsigma_t + \kappa_t + \theta_t, \qquad (19)$$

where

$$\phi_k = \chi_k; \quad \phi_z = \chi_z \rho_z; \quad \phi_g = \chi_g \rho_g;$$
$$\phi_r = \chi_r \rho_r; \quad \phi_w = \chi_w \rho_w; \quad \mu_t = \frac{\phi_z \varepsilon_t}{\rho_z};$$
$$\varsigma_t = \frac{\phi_g v_t}{\rho_g}; \quad \kappa_t = \frac{\phi_r u_t}{\rho_r}; \quad \theta_t = \frac{\phi_w s_t}{\rho_w}.$$

Rolling over Equation (19), and noticing that by construction $E_t\left(\mu_{t+1}\right) = 0$, $E_t\left(\varsigma_{t+1}\right) = 0$, $E_t\left(\kappa_{t+1}\right) = 0$ and $E_t\left(\theta_{t+1}\right) = 0$, we obtain:

$$E_t \hat{k}_{t+1} = \phi_k \hat{k}_t + \phi_z \hat{z}_t + \phi_g \hat{g}_t + \phi_r \hat{\tau}^r_t + \phi_w \hat{\tau}^w_t. \qquad (20)$$

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

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

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:

$$\phi_k = \frac{a_2}{1 - a_1\phi_k}, \qquad \phi_z = \frac{(a_1\phi_z + b_1)\rho_z}{1 - a_1\phi_k},$$
$$\phi_g = \frac{(a_1\phi_g + b_2)\rho_g}{1 - a_1\phi_k}, \qquad \phi_r = \frac{(a_1\phi_r + b_3)\rho_r}{1 - a_1\phi_k},$$
$$\phi_w = \frac{(a_1\phi_w + b_4)\rho_w}{1 - a_1\phi_k},$$

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.[6]

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

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

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 the driving force of the system. 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 expectatons case, conjecture (22) may be plugged in Equation (17a). Then, using Equations (17b), (17c), (17d) and (17e), and rearranging terms one gets the so-called 'actual law of motion' (ALM):

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

The linear system formed by the ALM – Equation (23) – and Equations (17b), (17c), (17d), (17e) and (17f) gives the adaptive learning solution to

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

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:

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

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

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 $(\phi_k, \phi_z, \phi_g, \phi_r, \phi_w)$. 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.[7] 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 centred 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, $(\hat{k}_t - \mathbf{x'_{t-1}} \mathbf{\Phi_{t-1}})$, corrected by the variance of the state variables, $\mathbf{R^{-1}_{t-1}} \mathbf{x_{t-1}}$, 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.[8]

---

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

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

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

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

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 funcion, α, the average growth rate of the U.S. GDP per capita, γ, and the 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 coffecient ρ_z is estimated to 0.85. Similarly, public expenditure shocks are the residuals from regressing an AR(1) process on the detrended share of public expenditure over GDP, which reduces to Equation (8). The autoregressive coffecient 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.[9] 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.[10]

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 behaviour 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 in detrended output, compared to 57% for the model with rational expectations. 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.[11] Thus, the model with learning accounts for 32% of the cumulative

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

[10] 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.

[11] As shown in Appendix C, the results mentioned in the text are independent of the inisialisation of the variance-covariance matrix $\mathbf{R}$. We used random generated data as

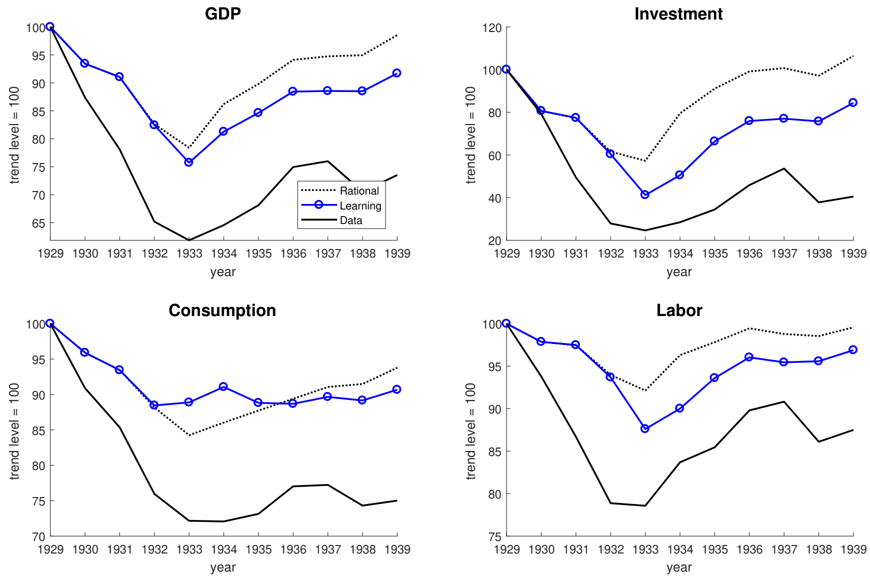

Figure 1: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with markers), data (black line).

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, 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 uncertainty 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,

benchmark, but the model with theoretical moments does just as good.

this figure represents a visual decomposition of the ALM – Equation (21) or (23), with rational or adaptive expectations, respectively – element by element. The graphs show that while with rational expectations the expec-

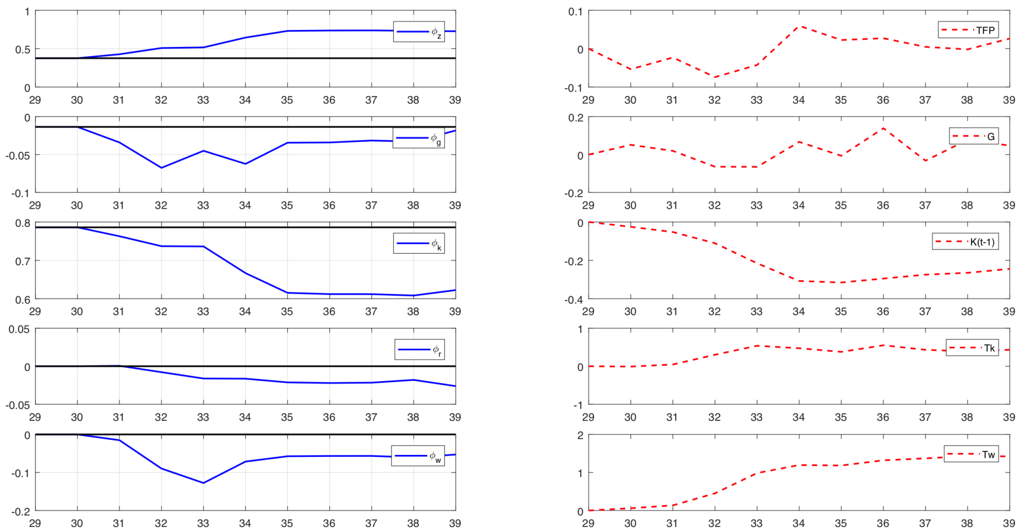

Figure 2: Simulations. Expectation coefficients (left column) and shocks (right column). Blue line: adaptive learning. Black line: rational expectations. Red line: data.

tation coefficients are obiously 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 agents'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 significanlty 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 uncertainty 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. 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.[12] When the economy is perturbated along more than one dimension, instead, forecast errors add to each other, thereby amplifying the distorting role of beliefs. Uncertainty matters more in complex, deep crises.

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 witnessess 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 unfavourable 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 the full structure of the model and are able to obtain

[12]Results available upon request.

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 is typically obtained by imposing an *ad hoc* PLM, meaning that there is no longer a learnable mapping between the rational expectation and the learning solution. Notice that this procedure may be criticised on the the ground that it gives the researcher some freedom in choosing the information set that agents are using when forming expectations in the model. On the other hand, when subject to random shocks, the model with this type of learning amplifies the volatility of economic variables and increase the persistence of fluctuations, while mantaining the same average behaviour over time as the model with rational expectations.[13]

In order to minimise the departure from rational expectations, we assume that households 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 coeficients associated to the PLM for the expectations of consumption and hours worked respectively for each state variable x.

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

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

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

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 and $\mathbf{\Phi_t}$ is the matrix of the coefficients corresponding to the PLMs and ζ_t stands for the speed of learning.

We initialise 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^j_i. We then plug in the measured shocks linked to the Great Depression and simulate the model.

[13]See Appendix D.

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 the 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 thorugh the law of motion of capital. As a result, here 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, induces a higher intertemporal substitution effect, with investment decreasing more than in the rational-expectations and adaptive-learning cases, and consumption, instead, increasing. As the drop in investment progressively affects capital, the capital-labour ratio eventually diminishes and remains significanlty below its counterpart in both the rational-expectations and adaptive models. The implied higher value of the real interest rate delays the recovery further. So, in this model, 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.[14]

[14]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)),

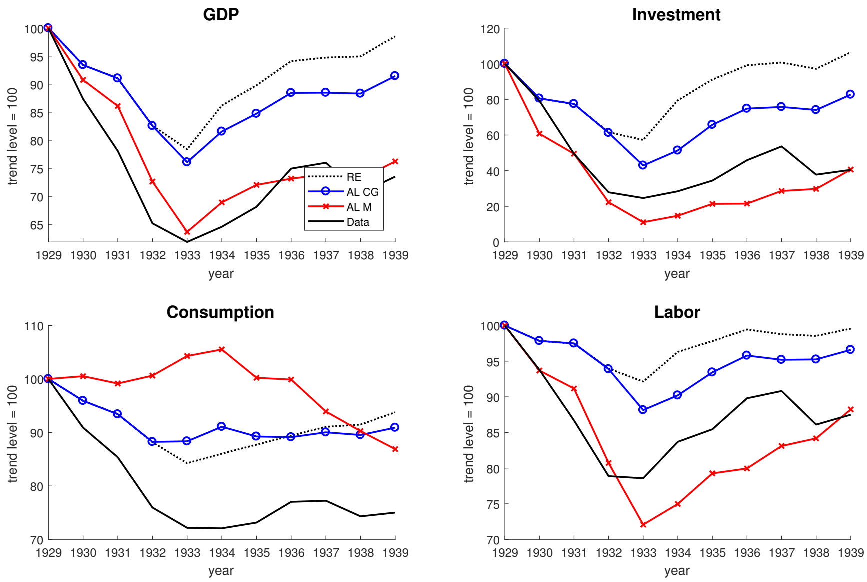

Figure 3: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with circles), model with learning about FOCs (red line with crosses), data (black line).

In Appendix E, 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.

which assumes away the wealth effect on labour supply. We found that high values of both the relative risk aversion ($\geq$ 2.5) and the Frisch elasticity of substitution ($\sim$ 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 restpect 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.

6 Conclusions

This article introduces adaptive learning in a stylised DGE 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 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 strenght of DGE modelling.

References

Aguilar, Pablo and Jesús Vázquez, "An estimated DSGE model with learning based on term structure information," *Macroeconomic Dynamics*, 2019, pp. 1–31.

_ **and** _ , "Adaptive learning with term structure information," *European Economic Review*, 2021, *134*, 103689.

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.

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 Friedman-Schwartz 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.

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.

Slobodyan, Sergey and Raf 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.

Weder, M., "The Role of Preference Shocks and Capital Utilization in the Great Depression," *International Economic Review*, 2006, *47*, 1247–1268.

A Computations

A.1 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:

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

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

$$\hat{y}_t = z_t + (1-\alpha)\hat{n}_t + \alpha\hat{k}_{t-1} \tag{30}$$

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 $\ln(\exp(z_t)) + \ln(1 - \alpha) - \alpha \ln(n_t) + \alpha \ln(k_{t-1})$, 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:

$$\hat{w}_t = z_t + \alpha(\hat{k}_{t-1} - \hat{n}_t) \tag{31}$$

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

$$\hat{r}_t = z_t + (1-\alpha)(\hat{n}_t - \hat{k}_{t-1}) \tag{32}$$

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:

$$g_{ss}\hat{g}_t = w_{ss}\tau_{ss}^w n_{ss}(\hat{w}_t + \hat{\tau}_t^w + \hat{n}_t) + r_{ss}\tau_{ss}^r k_{ss}(\hat{r}_t + \hat{\tau}_t^r + \hat{k}_{t-1}) - \frac{tr^{ss}\hat{tr}_t}{g_{ss}} \tag{33}$$

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:

$$y_{ss}\hat{y}_t = c_{ss}\hat{c}_t^p + i_{ss}\hat{i}_t + g_{ss}\hat{g}_t \tag{34}$$

The linearised capital law of motion (12):

$$\hat{i}_t = \frac{k^{ss}}{i^{ss}}\gamma\hat{k}_{t-1} - \frac{k^{ss}}{i^{ss}}(1-\delta)\hat{k}_{t-1} \tag{35}$$

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:

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

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

$$\frac{c_{ss}\hat{c}_t^p + g_{ss}\eta\hat{g}_t}{c_{ss}^p + g_{ss}} = \frac{c_{ss}\hat{c}_{t\,t+1}^p + g_{ss}\eta\hat{g}_{t+1}}{c_{ss}^p + g_{ss}} - \left(\frac{r_{ss}}{r_{ss}+1-\delta}\right)\left(\hat{z}_{t+1} + (1-\alpha)(\hat{n}_{t+1} - \hat{k}_t) - \tau^r_{ss}\hat{\tau}^r_{t+1}\right). \tag{36}$$

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:

$$\hat{n}_t = \left(\frac{1-n^{ss}}{n^{ss} + \alpha(1-n^{ss})}\right)\left(\hat{z}_t - \alpha\hat{k}_{t-1} - \left(\frac{c^{pss}\hat{c}_t^p + g^{ss}\eta\hat{g}_t}{c^{pss} + g^{ss}}\right) - \tau^{wss}\hat{\tau}_t^w\right) \tag{37}$$

A.2 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) 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 $n_1 = (1-\alpha)\left(\frac{1-n_{ss}}{n_{ss}+\alpha(1-n_{ss})}\right)$ This equation is going to be crucial in obtaining the reduced form of the model, for the sake of simplicity we rewrite it as:

$$\hat{c}_t^p = z_0\hat{z}_t + k_0\hat{k}_{t-1} - k_1\hat{k}_t - g_0\hat{g}_t - \tau_0^w\hat{\tau}_t^w, \tag{38}$$

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

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

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:

$$c_0\hat{c}^p_t = (1+r_1)c_0\hat{c}^p_{t+1} + (\rho_g - 1 + r_1\rho_g)g_1\hat{g}_t - (r_0+r_1)\rho_z\hat{z}_t + (r_0(1-\alpha)-r_1)\hat{k}_t + r_1\rho_w\tau^{wss}\hat{\tau}^w_t + r_0\rho_r\tau^{rss}\hat{\tau}^r_t,$$

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:

$$\hat{k}_t = a_1E_t\hat{k}_{t+1} + a_2\hat{k}_{t-1} + b_1\hat{z}_t + b_2\hat{g}_t + b_3\hat{\tau}^r_t + b_4\hat{\tau}^w_t, \tag{39}$$

where $a_0 = -c_0k_1 - c_1k_0 - r_0(1-\alpha) + r_1\alpha$, $c_1 = (1+r_1)c_0$,

$$a_1 = -\frac{c_1k_1}{a_0}, a_2 = -\frac{c_0k_0}{a_0}, b_1 = -\frac{c_0z_0 + c_1\rho_zz_0 - (r_0+r_1)\rho_z}{a_0},$$

$$b_2 = c_0g_0 - \rho_gc_1g_0 + (\rho_g - 1 + r_1\rho_g)g_1/a_0,$$

$$b_3 = \frac{r_0\tau^{rss}\rho_r}{a_0}, \text{ and } b_4 = \frac{c_0\tau^w_0 - c_1\tau^w_0\rho_w + r_1\tau^{wss}\rho_w}{a_0}.$$

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.

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

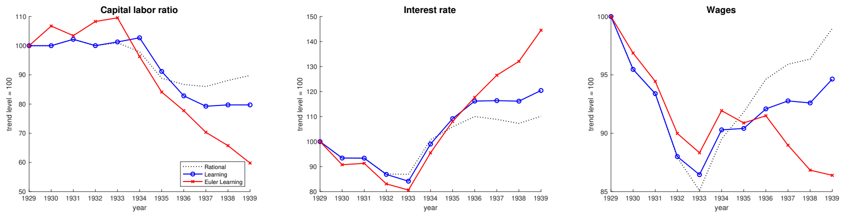

Figure 4: Simulations. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with circles), model with learning about FOCs (red line with crosses).

C 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.

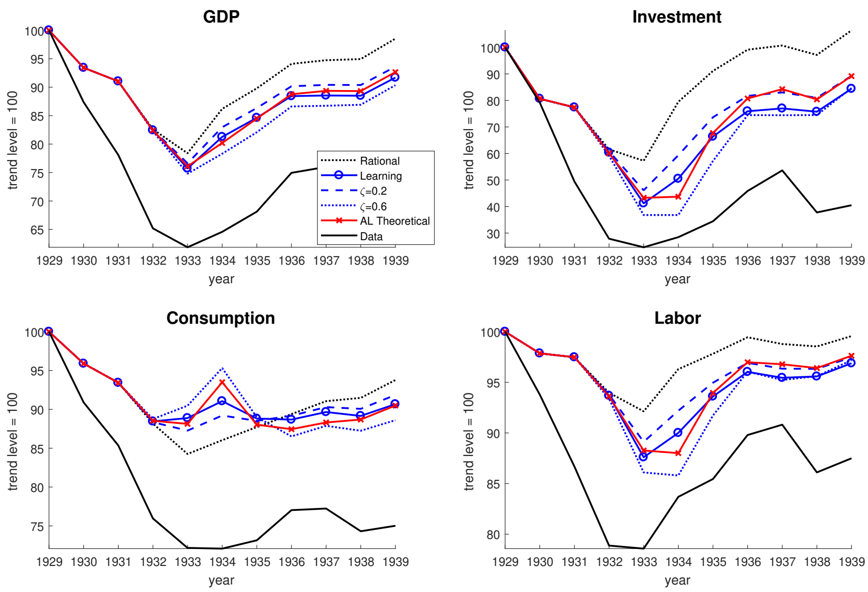

Figure 5: Simulations under different gains. Model with rational expectations (black-dotted line), model with adaptive learning (blue line with cercles) and different gains (dashed lines) and recursive least squares (red line with crosses).

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

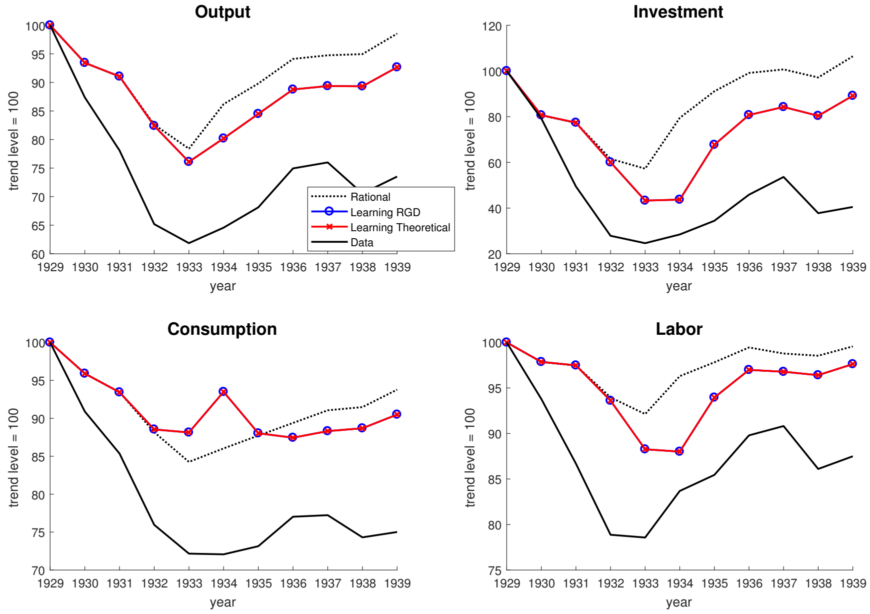

Figure 6: Simulations of the learning model under different intial conditions mechanisms. Model with rational expectations (black-dotted line), learning model initialized with Randomly Generated Data (blue line with cercles) and learning model initialized with the model implied theoretical momments (red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

model.[15]. 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.

D 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 7 shows that the model with adaptive learning and constant gain converges quite fastly to the rational expectation solution and shows similar volatility, whereas

[15]For the specific representation of this matrix, see Carceles-Poveda and Giannitsarou (2007)

the model with FOC-learning à la Milani (Euler learning) tends to amplify the volatility, especially in reaction to large shocks.

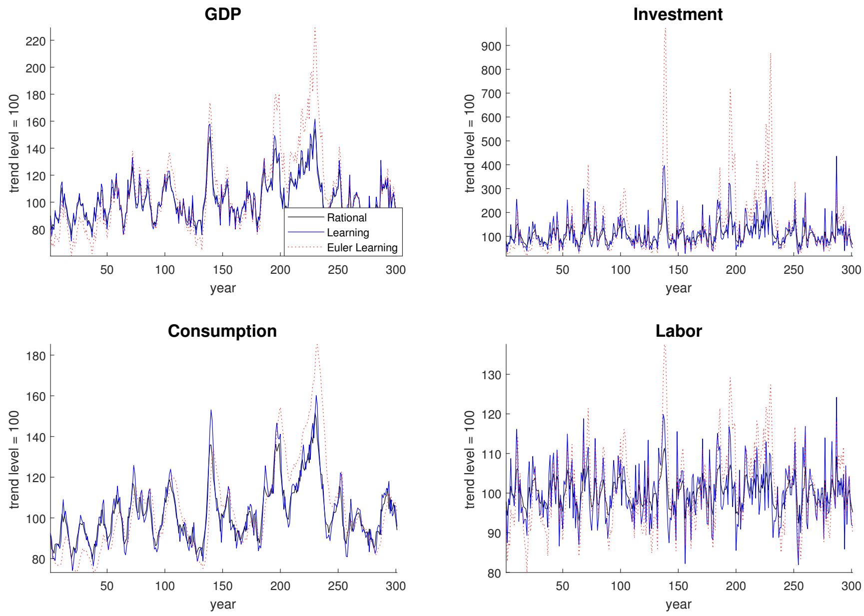

Figure 7: Model dynamics under different expectations, measured in deviations from the steady-state. Model with rational expectations (black line), model with adaptive learning (blue line) and Euler learning (red-dotted line).

E Alternative ways of modelling learning

E.1 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 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 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 8 shows the difference in the dynamics of output and consumption for our excercise in Section 4 as the share of rational agents shrinks.

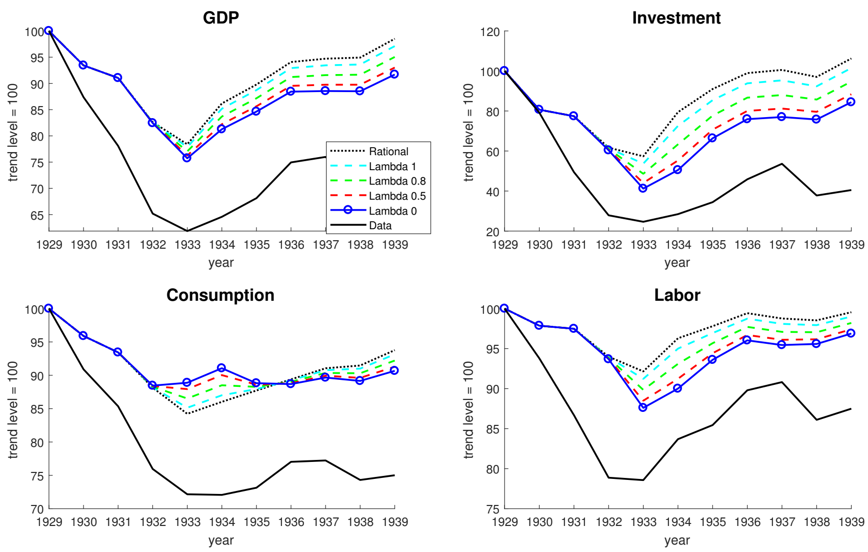

Figure 8: Simulations under different degrees of hybrid expectations. From 1, black dotted line(purely rational) to 0, blue line with cercles (learning).

E.2 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 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

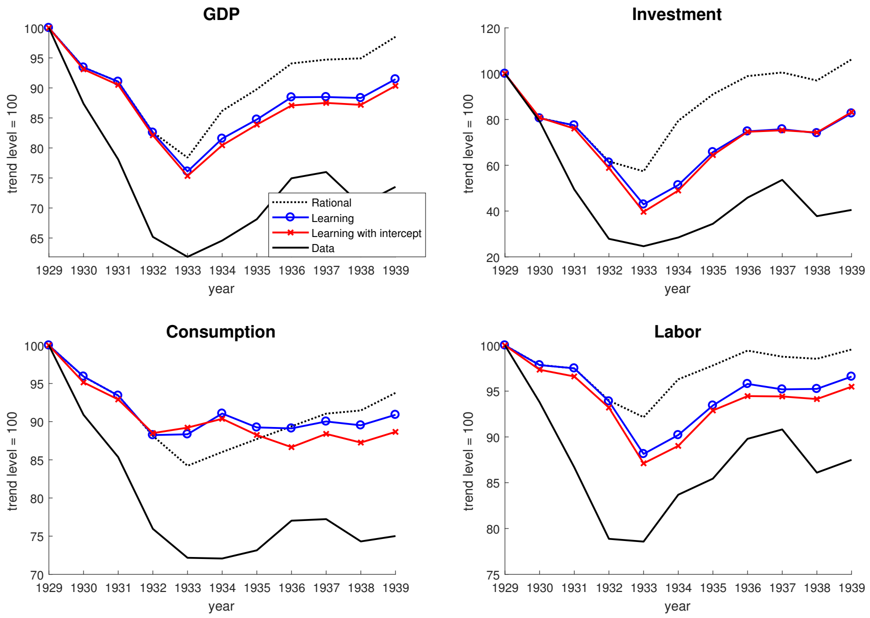

Figure 9: Simulation. Model with rational expectations (black-dotted line), baseline learning model (blue line with cercles) and learning model with an intercept (red line with crosses).

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:

$$\mathbf{\Phi_t} = \mathbf{\Phi_{t-1}} + \zeta_t \mathbf{R}_{\mathbf{t-1}}^{-1}\mathbf{x_{t-1}}(\hat{k}_t - \mathbf{x}'_{\mathbf{t-1}}\mathbf{\Phi_{t-1}}), \tag{40}$$

$$\mathbf{R_t} = \mathbf{R_{t-1}} + \zeta_t(\mathbf{x_{t-1}}\mathbf{x}'_{\mathbf{t-1}} - \mathbf{R_{t-1}}), \tag{41}$$

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 9 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.

INSTITUT DE RECHERCHE
ÉCONOMIQUES ET SOCIALES

Place Montesquieu 3

1348 Louvain-la-Neuve

ISSN 1379-244X D/2022/3082/04

UCLouvain

IRES | Louvain Institute of Data Analysis and Modeling in economics and statistics