Endogenously (Non-)Ricardian Beliefs*

William A. Branch
University of California, Irvine

Emanuel Gasteiger
TU Wien

December 12, 2023

Abstract

This paper revisits the impact of fiscal shocks on inflation dynamics. Traditionally, such relationships have been attributed to exogenous shifts in policy regimes. In contrast, we focus on a novel belief formation mechanism at work – specifically, when Ricardian equivalence, or its failure, can arise endogenously. Agents choose between two forecasting models: in one, Ricardian beliefs emerge naturally; in the other, beliefs are non-Ricardian. A predictor selection process guides these choices. Employing least squares learning in an estimated New Keynesian model highlights the significant influence of fiscal shocks in the U.S. within non-Ricardian belief regimes. However, learning dynamics also initiate self-fulfilling transitions toward a Ricardian regime. Non-Ricardian beliefs generate significant wealth effects, particularly during the high inflation period of the 1970s and 2021-22.

JEL Classification: D82; D83; E40; E50

Keywords: adaptive learning, misspecification, escapes, fiscal theory of price level.

*Gasteiger is also affiliated with ISCTE-IUL. Financial support from the Austrian National Bank (OeNB) grant no. 18611, Fundação para a Ciência e a Tecnologia grant no. UIDB/00315/2020, the Berlin Economics Research Associates (BERA) program and the Fritz Thyssen Foundation is gratefully acknowledged. We thank seminar participants at University of Alberta, California State University Fullerton, U.C. Irvine, Universität Potsdam, Freie Universität Berlin, Humboldt-Universität zu Berlin, Tinbergen Institute, TU Wien, Universität Wien, the 'North American Econometric Society Summer Meeting 2021', the 'Workshop on Expectations in Dynamic Macroeconomic Models', the 'CeNDEF@20 Workshop' and the 'Bamberg Behavioral Macroeconomics Workshop' for many helpful comments.

1 Introduction

This paper offers a new perspective on the role of fiscal shocks in shaping inflation dynamics by investigating the influence of endogenous changes in beliefs. Much of the existing literature examines policy regime changes or explicit assumptions on beliefs, that violate Ricardian equivalence. We develop a model instead that highlights the importance of evolving beliefs as a key driver of departures from Ricardian equivalence: with the exception of an edge case, households typically exhibit non-Ricardian beliefs that impact consumption and, in turn, inflation. In doing so, we provide new insights into the complex relationship between fiscal shocks, inflation, and the beliefs held by economic agents. Our main finding, both theoretical and empirical, is as follows: evolving beliefs on their own generate fluctuations between non-Ricardian and Ricardian regimes.

Our approach adopts a "restricted perceptions" framework, in which economic agents possess imperfect knowledge of the economy and its dynamics (see, Evans and Honkapohja, 2001; Branch and McGough, 2018; Woodford, 2013). This framework allows us to explore the implications of agents' imperfect knowledge on inflation dynamics and the emergence of beliefs that do not necessarily correspond to the true state of the world. By focusing on the role of endogenous beliefs in driving inflation patterns, we contribute to a deeper understanding of the mechanisms through which fiscal shocks impact inflation within a conventional policy stance of a monetary authority focused on price stability and a fiscal authority adhering to long-run government solvency.

In our analysis, we utilize a standard New Keynesian framework as a laboratory for our ideas. The starting point builds on Evans et al. (2012), Eusepi and Preston (2018), and Woodford (2013), who show that Ricardian Equivalence can fail in models where agents have imperfect knowledge about aggregate variables and whether the path of future primary budget surpluses is sufficient to satisfy the government's intertemporal budget constraint. In these papers, imperfect knowledge and non-Ricardian beliefs arise because least-squares learning generates temporary fluctuations around the rational expectations equilibrium. In contrast, our agents form subjective expectations about payoff-relevant aggregate variables based on parsimonious forecasting models and update these expectations using least-squares learning. We allow agents to choose between two parsimonious forecasting models – one that includes the stock of debt and is consistent with Ricardian beliefs, and another that includes the primary budget surplus but is inconsistent with Ricardian beliefs. Agents select the model that performs best in terms of mean-squared error.

A theory of predictor choice between parsimonious models achieves two aims. First, it captures that rational expectations may be prohibitively expensive in the presence of significant computational and cognitive limitations. Still, with least-squares learning and endogenous model choice, beliefs satisfy a set of cross-equation restrictions, a salient feature of rational expectations. Second, the debt-based forecast model is a close approximation to rational expectations when all agents forecast with it. Thus, we do not *a priori* rule out Ricardian beliefs. In fact, estimating our model on U.S. data suggests that a non-Ricardian equilibrium is a stable limit point to agents' learning and model selection. Occasionally, though, the economy escapes the non-Ricardian equilibrium and Ricardian beliefs emerge for a stretch of time.

We define a *Ricardian wedge* as the distance between aggregate expectations and full information rational expectations. The Ricardian wedge is determined by two state variables: the distribution of agents across forecasting models and the coefficients within each model. We highlight the fragility of Ricardian Equivalence, showing that a necessary and sufficient condition for its validity is that all agents in the economy use the same Ricardian model for forecasting. The presence of even an infinitesimal fraction using an alternative model renders everyone's beliefs non-Ricardian. This fragility suggests that a non-Ricardian equilibrium is a plausible outcome of the learning process. However, least-squares learning can also generate endogenous fluctuations between Ricardian and non-Ricardian regimes.

Our empirical investigation, conducted using Bayesian estimation techniques, reveals the presence of non-Ricardian beliefs in the U.S. economy, with the extent of such beliefs fluctuating over time. Notably, we find that inflationary periods during the late 1960s and 1970s were driven by a growing prevalence of non-Ricardian beliefs, while disinflation and subsequent low inflation periods resulted from temporary shifts to a Ricardian regime. This finding suggests that the key question is not why fiscal shocks matter at certain times, but rather why inflation is not always a fiscal phenomenon.

Drawing parallels to the "Conquest" model by Sargent (1999), Cho et al. (2002), and Sargent et al. (2006) our study also underscores the pivotal role of evolving beliefs and regime shifts in inflation dynamics. These earlier models demonstrate how a central bank's time consistency problem combined with occasional shocks that alter the policymaker's view on the inflation-unemployment trade-off, can temporarily reduce inflation to its efficient rate. Our findings suggest a similar dynamic, though with emphasis on private sector expectations: shifts towards a Ricardian regime, albeit temporary, can lead to disinflation.

The empirical analysis confirms the prevalence of non-Ricardian beliefs within the U.S. economy. Specifically, the high inflation of the 1970s, and its resurgence in 2021-22 can be traced to the economy transitioning towards a non-Ricardian belief equilibrium. This pattern reaffirms the inherent fragility of Ricardian equilibria and the importance of endogenously non-Ricardian beliefs.

The remainder of the paper is organized as follows: Section 2 presents the economic environment and the theory of restricted perceptions, along with Bayesian model estimates and comparisons. Section 3 offers theoretical insights that shed light on the fragility of Ricardian equilibria and the role of least-squares learning in reaching non-Ricardian limit points. Section 4 discusses the main empirical findings. Finally, sections 5 and 6 review the related literature and provide concluding remarks.

2 Model: specification and estimation

We begin by describing the economic environment, beliefs, and equilibrium concept. Then, we estimate the model's parameters using Bayesian methods.

2.1 Woodford's (2013) model

The setting is a New Keynesian model, based on Woodford (2013), where households and firms have subjective beliefs about payoff-relevant aggregate variables.[1] Given these beliefs, households choose consumption, leisure, and one-period government debt, the only asset available to households, to solve their intertemporal optimization problem. In Woodford's (2013) framework, households turn over wage-setting and labor supply decisions to a union and are obligated to supply labor to a firm on the union's terms. Households also receive a lump-sum transfer of their share in firm profits.[2] This is a stylized assumption that renders the household's consumption rule

[1] The model environment is a simplified version of Eusepi and Preston (2018) where there are two assets, one-period government bonds in zero net-supply and longer maturity bonds. Eusepi and Preston (2018) demonstrate the role that maturity structure, combined with imperfect knowledge and learning, can play in generating non-Ricardian wealth effects.

[2] The shares in firms are illiquid, which makes government debt the only storable good. Eusepi and Preston (2018) show that this assumption is consequential for non-Ricardian beliefs. Though we abstract from these issues, it is worth bearing in mind that the issue is relevant within our non-Ricardian equilibrium.

analogous to the one in a model where the household receives a stochastic endowment. However, because firms are monopolistically competitive and face a Calvo (1983) nominal pricing friction, there is endogenous variation in hours and output. All exogenous shocks follow stationary AR(1) processes.

HOUSEHOLDS. Woodford (2013) derives an individual's consumption function,

$$\begin{aligned} c_t^i = (1-\beta)b_t^i + \sum_{T=t}^{\infty} \beta^{T-t} E_t^i \{ &(1-\beta)(Y_T - \tau_T) - \beta\sigma(\beta i_T - \pi_{T+1}) \\ &+ (1-\beta)s_b(\beta i_T - \pi_T) - \beta(\bar{c}_{T+1} - \bar{c}_T) \}, \end{aligned} \tag{1}$$

or, written recursively as

$$c_t^i = (1-\beta)\left[b_t^i + (Y_t - \tau_t) - s_b \pi_t\right] - \beta[\sigma - (1-\beta)s_b] i_t + \beta \bar{c}_t + \beta E_t^i v_{t+1}^i, \tag{2}$$

where v_t^i is a subjective composite variable that comprises all payoff-relevant aggregate variables over which a household formulates subjective beliefs:

$$v_t^i = (1-\beta)(Y_t - \tau_t) - [\sigma - (1-\beta)s_b](\beta i_t - \pi_t) - (1-\beta)\bar{c}_t + \beta E_t^i v_{t+1}^i.$$

The variables, written as log-deviations from steady-state, $b_t^i, Y_t, \pi_t, i_t, \tau_t, \bar{c}_t$ are, respectively, the individual's holdings of real government debt, aggregate output, the inflation rate, the nominal interest rate, lump-sum taxes, and a preference shock. The government uses lump-sum taxes and debt to finance its consumption of an exogenous sequence G_t. The parameter $0 < \beta < 1$ is the discount rate, σ is the elasticity of intertemporal substitution, and $s_b \equiv \bar{b}/\bar{Y}$ is the steady-state debt-to-GDP ratio. The fiscal policy instrument is the real primary surplus $s_t \equiv \tau_t - G_t$.

The consumption function takes a standard permanent income formulation relating consumption to the annuitized present-value of financial and non-financial income.[3] The first two terms in (1) dictate how consumption responds to government bond holdings and disposable income, respectively. The first term is sometimes called a "wealth effect". The third term, parameterized by σ, captures an intertemporal substitution effect resulting from variations in the (perceived) *ex-ante* real interest rate. The fourth term, pre-multiplied by s_b, is the perceived real return on government bond holdings.

[3]Derivation of (1) is entirely standard. We refer readers to Woodford (2013) for details.

Woodford (2013) describes this term as an "income effect." Note from the final term that a positive preference shock, $\bar{c}_t$, implies a stronger desire for contemporaneous consumption.

Following Eusepi and Preston (2018), Woodford (2013) derives equation (1) without assuming that individuals have structural knowledge about the government's intertemporal budget constraint. Even though there is passive fiscal policy, individuals do not necessarily know this or the other structural features of the economy, so that they may have imperfect knowledge about the structural form of the government's endogenously determined budget constraint. Instead, they form subjective beliefs over the evolution of aggregate variables.[4] If they get those beliefs right, they will properly account for the evolution of debt, and their beliefs will be Ricardian. Otherwise, beliefs may be non-Ricardian.

Ricardian beliefs arise with the following condition on *beliefs*

$$E_t^i \left\{ \sum_{T=t}^{\infty} \beta^{T-t} \left[s_T - s_b(\beta i_T - \pi_T) \right] \right\} = b_t. \tag{3}$$

By imposing Ricardian beliefs onto the consumption rule (1), (3) directly imposes that the household properly forecasts the path for future surpluses. Without assuming (3), there are potentially non-Ricardian effects in current bond holdings ("wealth effect") and the perceived present value of the future real returns on bonds ("income effect"). Ricardian beliefs, therefore, lead to a consumption rule that depends only on the household's subjective beliefs about future paths for disposable income and real interest rates. Conversely, by not *a priori* imposing Ricardian beliefs, households may perceive their current bond holdings and anticipated future real returns as real wealth, and a change in the expected path for future surpluses can have a real effect on consumption. See Appendix 6 or Woodford (2013) for details.[5]

Firms. Monopolistically competitive firms face a nominal pricing friction based on

[4]On the surface, formulating expectations over future v_t^i seems to be adopting the Euler equation approach of one-step ahead forecasting and decision-making. However, the derivation of the consumption function and v_t^i is based on the infinite-horizon approach where the household's consumption/savings decisions solve their entire sequence of Euler equations, flow budget constraints, and transversality condition given their subjective beliefs. We show below how these consumption rules can be aggregated with heterogeneous agents.

[5]In all of the analysis below, the fiscal rule is *ex post* Ricardian, i.e., real primary surpluses will satisfy the government's intertemporal constraint. However, out of equilibrium, non-Ricardian beliefs could be consistent with explosive debt. The consequences of this, and its implications for strategic behavior, is an old issue in the fiscal theory of the price level literature (cf., Bassetto, 2002).

Calvo (1983). An individual firm j produces a differentiated good. With probability $0 < \alpha < 1$, it will adjust its previous price by the long-run target rate of inflation, assumed to be zero, and with probability $1-\alpha$, a firm receives an idiosyncratic signal to (optimally) reset the price. A firm j that can optimally reset price $p_t^*(j)$, relative to the previous aggregate price level p_{t-1}, will do so to satisfy the log-linear approximated first-order condition, written recursively,

$$p_t^*(j) = (1-\alpha\beta)\left(E_t^j p_t^{\text{opt}} - p_{t-1}\right) + (\alpha\beta)E_t^j p_{t+1}^*(j) + (\alpha\beta)\pi_t,$$

where $E_t^j p_T^{\text{opt}}$ is the perceived optimal price in T. The aggregate inflation dynamics are

$$\pi_t = (1-\alpha)p_t^*, \qquad \text{where} \qquad p_t^* \equiv \int p_t^*(j)dj. \tag{4}$$

Policy. Monetary policy is described by a Taylor (1993) rule,

$$i_t = \phi_\pi \pi_t + \phi_y y_t + w_t, \tag{5}$$

where the monetary policy shock follows $w_t = \rho_w w_{t-1} + \varepsilon_{wt}$, $\varepsilon_{wt} \sim N(0, \sigma_w^2)$.

A Leeper (1991) rule for the real primary surplus characterizes fiscal policy:

$$s_t = \phi_b b_t + z_t, \tag{6}$$

where the surplus shock is $z_t = \rho_z z_{t-1} + \varepsilon_{zt}$, $\varepsilon_{zt} \sim N(0, \sigma_z^2)$. The government also faces a flow budget constraint

$$b_{t+1} = \beta^{-1}[b_t - s_b \pi_t - s_t] + s_b i_t. \tag{7}$$

The steady-state debt-to-GDP ratio s_b plays a role in the results presented below. When $s_b = 0$, the bond and primary surplus paths are exogenous, while $s_b > 0$ implies that they are endogenous and affected, in part, by monetary policy.[6]

Throughout, the analysis focuses on the *active* monetary and *passive* fiscal policy

[6]This formulation arises in a cashless environment that allows us to abstract from the effect of monetary aggregates appearing in the consolidated budget constraint.

regime:

$$1 < \phi_\pi + \frac{1-\beta}{\kappa}\phi_y$$
$$(1-\beta) < \phi_b < 1.$$

Under the benchmark *rational expectations hypothesis*, there is local determinacy (see, Leeper, 1991) implying that this locally unique *rational expectations equilibrium* displays Ricardian equivalence and is stable under least-squares learning (see, Evans and Honkapohja, 2007).

2.2 Temporary equilibrium with heterogeneous beliefs

The income-expenditure identity is given by

$$Y_t = \int c_t^i di + G_t. \tag{8}$$

Combining (2) and (8) with the bond-market clearing condition $b_t \equiv \int b_t^i di$, computing $v_t \equiv \int v_t^i di$, and averaging over expectations, allows us to express aggregate demand as the "IS equation" without *a priori* imposing Ricardian beliefs:

$$y_t = g_t - \sigma i_t + (1-\beta)b_{t+1} + \hat{E}_t v_{t+1}, \tag{9}$$

where $y_t \equiv Y_t - Y_t^n$, $g_t \equiv \bar{c}_t + G_t - Y_t^n$ is a composite exogenous disturbance, such that $g_t = \rho_g g_{t-1} + \varepsilon_{gt}$, $\varepsilon_{gt} \sim N(0, \sigma_g^2)$. The aggregate expectations operator $\hat{E}$ is defined as $\hat{E}_t(x) = \int E_t^i(x)\, di$, for any variable x.

Given that heterogeneous beliefs lead to a non-degenerate cross-sectional wealth and consumption distribution, some readers may be surprised that individual household bond holdings do not appear in the aggregate demand equation. However, this results from several simplifying assumptions in Woodford (2013). First, assumptions about the labor market and the distribution of firm profits imply that future non-financial income is a proportion of aggregate output beyond the agent's control. So, household consumption decisions depend on expectations about variables beyond their control. Second, in this setting, a temporary equilibrium path consists of local perturbations around a non-stochastic steady-state in which all agents hold identical beliefs. In this sense, households' beliefs are not *too* heterogeneous. Finally, household debt holdings enter linearly in the approximated economy, and, as a result, individual bond holdings do not matter for the aggregate output path.[7]

On the firm side, applying the law of iterated expectations and aggregating across all firms results in an aggregate New Keynesian Phillips Curve:

$$\pi_t = (1-\alpha)\,\beta \hat{E}_t p^*_{t+1} + \kappa y_t + u_t,$$

where $\kappa \equiv [(1-\alpha)(1-\alpha\beta)\zeta]/\alpha$, and the cost-push shock is $u_t \equiv \{[(1-\alpha)(1-\alpha\beta)]/\alpha\}\mu_t$, with $u_t = \rho_u u_{t-1} + \varepsilon_{ut}$, $\varepsilon_{ut} \sim N(0, \sigma_u^2)$.

We now define a *temporary equilibrium* for this economy.

Definition 1 *Given a distribution of beliefs* $\left(E^i_t v_{t+1}, E^i_t p^*_{t+1}\right)_i$ *a temporary equilibrium is a triple* (b_{t+1}, π_t, y_t) *and a policy* (s_t, i_t) *so that the bond and goods markets clear and the government budget constraint is satisfied. In particular, the following equations are satisfied*

$$\begin{aligned}
b_{t+1} &= \beta^{-1}\left[b_t - s_b\pi_t - s_t\right] + s_b i_t \\
\pi_t &= (1-\alpha)\beta\hat{E}_t p^*_{t+1} + \kappa y_t + u_t \\
y_t &= g_t - \sigma i_t + (1-\beta)b_{t+1} + \hat{E}_t v_{t+1} \\
v_t &= (1-\beta)\left(b_{t+1} - b_t\right) - \sigma\left(i_t - \pi_t\right) + \hat{E}_t v_{t+1}.
\end{aligned}$$

2.3 Model misspecification

This section details expectation formation.

2.3.1 A restricted perceptions approach

Under full-information rational expectations, the equilibrium law of motion takes the form

$$\begin{bmatrix} \pi_t \\ v_t \\ y_t \end{bmatrix} = A \begin{bmatrix} b_t \\ s_t \end{bmatrix} + \eta_t,$$

where η_t is a vector of composite disturbances and A is conformable. It follows that to formulate rational expectations, the agents adopt linear forecast rules that depend on the stock of beginning-of-period debt, b_t and the primary surplus, s_t.

[7] An extension to a setting where heterogeneous expectations give rise to a non-trivial aggregate role to the cross-sectional wealth distribution is potentially important but beyond the scope of the present study.

In our framework, agents navigate their information landscape, optimizing their statistical forecasts based on their available information and abilities. Drawing from the econometric learning literature and the "cognitive consistency principle" as outlined by Evans and Honkapohja (2001), we model agents as skilled economists: they favor well-specified yet parsimonious econometric models. Parsimony is an appropriate response to the complexity of economic forecasting and the inherent limitations on degrees of freedom. Moreover, as Andre et al. (2022) show, forecasters often adopt narratives that guide their selective use of information, i.e., they do not use all information even when it is available. Our model captures this dynamic, allowing for the endogenous evolution of these narratives over time.

Explicitly, we assume that computational and cognitive limitations render full information rational expectations prohibitively expensive.[8] Consequently, we assume that agents forecast using one of two parsimonious models, each incorporating a single fiscal variable – either s_t or b_t. This approach not only enables a straightforward formalization of endogenously (non-)Ricardian beliefs but also aligns with empirical studies that utilize a limited number of fiscal indicators (Favero and Giavazzi, 2012).

By adopting this parsimonious modeling approach, we facilitate the emergence of a restricted perceptions equilibrium (RPE) wherein agents' expectations come from the linear projection of aggregate variables onto the space spanned by their parsimonious models. In a misspecification equilibrium (ME), the population distribution across the two forecasting models is determined endogenously by agents' discrete choices. Combined, the equilibrium distribution allows for the possibility of Ricardian equivalence emerging endogenously from agents' model choices.

While our equilibrium concept relaxes the model consistency of rational expectations, it maintains many cross-equation restrictions characteristic of rational expectations models. Our approach is particularly suited to the study of fiscal shocks and inflation, given its ability to capture endogenous regime shifts and its roots in the long tradition of empirical research using a limited set of fiscal indicators.

Admittedly, the restriction on regressors in agents' econometric models is somewhat ad hoc. However, alternative specifications for misspecified beliefs would require agents to possess extensive knowledge about the structural features of the economy

[8] It would be straightforward, but very costly to estimate, an extension that includes a third choice which consists of all variables but at a cost. In equilibrium the three models would produce mean-square forecast errors of a similar magnitude. Thus, for a sufficiently large cost there would continue to be a distribution across all models and each model, including the correctly specified model, would be non-Ricardian. This latter fact is a consequence of the model's self-referentiality.

– knowledge that extends beyond their models and includes the beliefs, constraints, and decision rules of other agents, such as the government's commitment to satisfying solvency constraints. Our approach, by contrast, offers a more natural and tractable way to model the influence of endogenous beliefs on inflation dynamics.

2.3.2 Equilibrium

Expectations come from one of the following forecasting models, or, perceived laws of motion (PLM):

$$\begin{aligned} PLM_s: \ \mathbf{Z}_t = \boldsymbol{\psi}^{s\prime} X^s_{t-1} + \eta_t \Rightarrow E^s_t \mathbf{Z}_{t+1} = \boldsymbol{\psi}^{s\prime} X^s_t \\ PLM_b: \ \mathbf{Z}_t = \boldsymbol{\psi}^{b\prime} X^b_{t-1} + \eta_t \Rightarrow E^b_t \mathbf{Z}_{t+1} = \boldsymbol{\psi}^{b\prime} X^b_t, \end{aligned}$$

where $\mathbf{Z}'_t = (v_t, p^*_t, b_{t+1})$, $X^s_t = (s_t, g_t, u_t, w_t, z_t)'$, $X^b_t = (b_t, g_t, u_t, w_t, z_t)'$, η_t is a perceived noise, and the coefficient matrix, for $k = \{s, b\}$,

$$\boldsymbol{\psi}^k = \left(\psi^k, \Gamma^k\right)',$$

$\psi^k = \left(\psi^k_v, \psi^k_p\right)'$ and Γ^k is the coefficient for b_{t+1}.[9] In a *restricted perceptions equilibrium* (RPE) the coefficients will satisfy the least-squares orthogonality condition:

$$EX^k_{t-1}\left(\mathbf{Z}_t - \boldsymbol{\psi}^{\boldsymbol{k}\prime} X^k_{t-1}\right)' = 0.$$

Beliefs, parameterized by $\boldsymbol{\psi}^k$, are derived from the optimal projection of the aggregate variables $\mathbf{Z}_t$ onto the restricted explanatory variable x^k_t. It follows that

$$\boldsymbol{\psi}^k = \left[EX^k_{t-1} X^{k\prime}_{t-1}\right]^{-1} EX^k_{t-1}\mathbf{Z}'_t \equiv S\left(\boldsymbol{\psi}^k\right).$$

Definition 2 *A* **restricted perceptions equilibrium** *is a fixed point* $\boldsymbol{\psi}^k_* = S\left(\boldsymbol{\psi}^k_*\right)$.

We do not impose *a priori* which of the PLM's individuals and firms use to form expectations. Instead, we confront them with a discrete choice: they can forecast by

[9] A brief remark about a timing assumption. Here, we follow Woodford (2013), in assuming that agents project the state variables onto the *lagged* regressors. We could alternatively assume that they regress the state onto *contemporaneous* regressors and it would not greatly impact the equilibrium results. However, the timing convention followed here has two benefits. First, it simplifies many of the analytic expressions. Second, in the quantitative analysis below, we implement a real-time learning version of the model and the timing avoids a potential multicollinearity problem.

including s_t ("model-s") or b_t ("model-b"), and like the selection of model parameters, they will do so to minimize their forecast errors. We adopt the rationally heterogeneous expectations approach first pioneered by Brock and Hommes (1997), extended to stochastic environments by Branch and Evans (2006). Agents make a predictor selection in a random-utility setting, and the agents will only select the best-performing statistical models in the limit of vanishingly small noise.

Let n denote the fraction of agents selecting model-s, leaving $1-n$ of the population forecasting with model-b.[10] They rank these choices by calculating the relative mean square error (MSE):

$$EU^k = -E\left[(\mathbf{Z}_t - E_t^k[\mathbf{Z}_t^k])\right]' \times \mathbf{W} \times E\left[(\mathbf{Z}_t - E_t^k[\mathbf{Z}_t^k])\right], \qquad k = \{s, b\}, \tag{10}$$

where $\mathbf{W}$ is a weighting matrix.[11] We define relative predictor performance $F(n)$: $[0,1] \to \mathbb{R}$ as $F(n) \equiv EU^s - EU^b$.

The distribution of agents across the two forecasting models, n, is pinned down according to the multinomial logit (MNL) map (see, e.g., Branch and Evans, 2006)

$$n = \frac{1}{2}\left\{\tanh\left[\frac{\omega}{2}F(n)\right] + 1\right\} \equiv T_\omega(n),$$

where ω denotes the "intensity of choice". The MNL map – also, the "T-map" – states that the fraction of agents adopting model-s, n, is an increasing function of its relative forecast accuracy, measured by the function $F(n)$.

Definition 3 *A* **misspecification equilibrium** *is a fixed point* $n_* = T_\omega(n_*)$.

An immediate consequence of the continuity of $T_\omega : [0,1] \to [0,1]$ is that there exists a misspecification equilibrium: see Appendix B for analytic details on existence. The neoclassical case $\omega \to \infty$ warrants special attention. In this case, agents only select the best-performing statistical models. It turns out that, in this case, one can learn quite a bit about the set of misspecification equilibria by studying the endpoints to $F(n)$. For instance, when $F(0) < 0, F(1) < 0$ – that is, the model-b forecasts best when all agents use model-b, or, if they all use model-s – then $n_* = 0$ is a misspecification equilibrium.

[10] For simplicity, we assume that households and firms are distributed across models identically, a simplification that could be relaxed. Instead, there could be a distribution n_h of households across models and a fraction n_f of firms. It would be straightforward to generalize this way at the cost of an expanded state vector.

[11] For simplicity, and without loss of generality, we set $W = I$.

Conversely, when $F(0) > 0$ and $F(1) > 0$, then $n_* = 1$ is a misspecification equilibrium. Outside of these polar cases, there is also the possibility of multiple misspecification equilibria, $n = \{0, \hat{n}, 1\}$, for some $0 < \hat{n} < 1$, that arises when $F(0) < 0, F(1) > 0$. As we will see, the $n = 0$ misspecification equilibrium is a self-confirming equilibrium with *weakly Ricardian beliefs*, and the $n = 1$ will correspond to homogenous *non-Ricardian beliefs.*[12]

The neoclassical limiting case, $\omega \to \infty$, helps build intuition, but a finite ω is relevant in practice. The multinomial approach, i.e., a finite ω, has a venerable history in discrete decision-making because it provides an elegant way of introducing randomness into discrete decision-making. Young (2004) shows that randomness in forecasting, much like mixed strategies in actions, provides robustness against model uncertainty and flexibility in self-referential economies. The intensity of choice parameter ω is inversely related to the idiosyncratic random utility innovation and, thereby, parameterizes model uncertainty. In particular, larger values of ω parameterize less model uncertainty with the neoclassical case $\omega \to \infty$ representing no uncertainty at all. Extensive literature tests for dynamic predictor selection using empirical MNL models and typically finds finite values for ω (c.f. , Branch, 2004). The parameter ω is an object in our estimation below.

2.3.3 Learning

The rational expectations hypothesis posits that subjective expectations do not have an independent role in the data-generating process, as cross-equation restrictions enforce consistency between agents' perceived and true laws of motion. A misspecification equilibrium also comprises a set of cross-equation restrictions that emerge through the least squares orthogonality condition. However, this paper focuses on the evolution of beliefs between (non-)Ricardian regimes, and thus, we develop a real-time learning model that converges to a misspecification equilibrium as its limit point.

Our learning model is a recursive adaptive algorithm with a constant gain, which assigns geometrically declining weights to past observations (see Marcet and Sargent, 1989; Evans and Honkapohja, 2001). This learning rule exhibits several notable features. First, it is a robust estimator in the presence of model misspecification and uncertainty (Evans et al., 2010), and it performs well econometrically when coefficients drift. Real-time learning and model selection generate both of these characteristics.

[12]For discussion of self-confirming equilibria see Sargent (1999). A self-confirming equilibrium is a stronger concept than RPE as it requires that agents' beliefs are correct, though they may be misspecified off the equilibrium path.

Second, the learning model extends the vector of state variables to include beliefs, enabling it to account plausibly for the persistence observed in U.S. macroeconomic data. Many modelers use medium-scale New Keynesian models that incorporate persistence mechanically through habits and price indexation. However, it has been shown (see Milani, 2007) that estimated New Keynesian models favor specifications where inertia arises endogenously through beliefs.

Finally, the real-time learning approach captures endogenous regime-switching beliefs in our empirical model while adhering to our restricted perceptions perspective. The key to interpreting the model estimates lies in understanding how learning can generate an "escape," as described by Sargent (1999), Cho et al. (2002), and Williams (2019). This feature allows our model to capture the dynamic nature of agents' beliefs as they respond to evolving macroeconomic conditions.

Extending the two restricted forecasting models to this more general environment, we can write

$$E_t^k x_{j,t+1} = \left(\psi_{j,t-1}^k\right)' X_{k,t-1},$$

where, for $j = \{v, p\}$ and $k = \{s, b\}$, $x_{j,t} \in \{v_t, p_t^*\}$, and $X'_{k,t-1} = (k_{t-1}, g_{t-1}, u_{t-1}, w_{t-1}, z_{t-1})$. The shocks follow uncorrelated stationary AR(1) processes with parameters ρ_j, σ_j and $\sigma_{ij} = 0$ for all $i, j \in \{g, u, w, z\}$. The econometric learning process is a recursive Bayesian model based on Evans and Honkapohja (2001):

$$\psi_{j,t}^k = \psi_{j,t-1}^k + \gamma_1 \Gamma X_{k,t-1}\left(x_{j,t} - \left(\psi_{j,t-1}^k\right)' X_{k,t-1}\right) \tag{11}$$

$$MSE_{j,t}^k = MSE_{j,t-1}^k + \gamma_2\left[\left(x_{j,t} - \left(\psi_{j,t-1}^k\right)' X_{k,t}\right)^2 - MSE_{j,t-1}^k\right] \tag{12}$$

$$EU_t^k = -MSE_{v,t}^k - MSE_{p,t}^k$$

$$n_t = \frac{1}{2}\left\{\tanh\left[\frac{\omega}{2}\left(EU_t^s - EU_t^b\right)\right] + 1\right\}.$$

Equation (11) is a generalized stochastic gradient algorithm that emerges from a Bayesian time-varying parameter model where the coefficients follow a random walk.[13] The parameter $0 < \gamma_1 < 1$ is the "constant gain" as it governs the responsiveness of parameter updating to recent forecast errors. The parameter Γ controls the direction of drift in the parameters. In our empirical application, we adopt the standard stochastic

[13] A constant gain least-squares algorithm is another commonly employed version of econometric learning. This algorithm looks similar but replaces the parameter Γ with a recursively estimated covariance matrix for the regressors. The latter substantially raises the computational cost of the empirical procedure.

gradient with $\Gamma = I$.[14] It is evident from the stochastic gradient algorithm that it nests the restricted perceptions equilibrium. Equation (12) is a simple recursive estimator of the mean-squared forecast errors that geometrically discounts at rate $(1-\gamma_2)$. The learning gains, γ_1, γ_2, are critical objects in the estimation as they control the relative speed of coefficient updating and model selection. If these estimated gain coefficients are significantly different, the dynamical system is a "fast-slow" system, a feature studied by the literature on large deviations (Dupuis and Ellis (1997)). The idea is that slow-moving variables can generate large fluctuations that occur infrequently but with high probability.

2.4 Empirical results

Our main interest is an empirical assessment of the role played by endogenously (non-)Ricardian beliefs. To that end, we first describe the Bayesian methods used to estimate the model's key parameters and assess model fit *vis a vis* a restriction to rational expectations or only Ricardian beliefs.

2.4.1 Empirical methodology

After plugging in the policy rules, expectations, and recursive updating equations for the learning rules, the model, in non-linear state space form, is:

$$\begin{aligned} X_t &= g\left(X_{t-1}, \Theta\right) + Q(X_{t-1}, \Theta)\nu_t \\ \mathbb{W}_t &= f(X_t, \upsilon_t), \end{aligned}$$

where the state vector is

$$X_t' = \left(b_{t+1}, \pi_t, y_t, v_t, s_t, g_t, u_t, w_t, z_t, n_t, MSE_{st}, MSE_{bt}, \operatorname{vec}\left(\psi_t^s\right), \operatorname{vec}\left(\psi_t^b\right)\right),$$

$\operatorname{vec}(\cdot)$ is the vectorization operator, the observation variables are

$$\mathbb{W}_t' = \left(y_t, \pi_t, s_t, b_{t+1}, i_t\right),$$

and the parameter vector is

$$\Theta' = \left(\kappa, \alpha, \phi_\pi, \phi_y, \phi_b, \rho_g, \rho_u, \rho_w, \rho_z, \sigma_g, \sigma_u, \sigma_w, \sigma_z, \omega, \gamma_1, \gamma_2\right).$$

[14] As in Sargent and Williams (2005), it is an avenue for future research to consider the role different priors about the direction of drift impacts time-varying Ricardian beliefs.

The measurement and state disturbances are υ_t, ν_t respectively. Our sample for the observed variables is 1955.1-2007.3.[15] We measure y_t as the log difference between output and the CBO's measure of potential output. We measure π_t from the PCE index. We compute b_t and s_t as the debt-GDP ratio and primary surplus-GDP ratio, respectively. All variables are deviations from the mean and annualized.

The empirical exercise aims to identify reasonable values for the parameters in Θ to explain the U.S. economy over the sample period. We are especially interested in inferences on the latent state variable, n_t, measuring the extent of (non-)Ricardian beliefs over the period. To this end, we adopt Bayesian methods and use a simulation-based technique called the particle filter to approximate the likelihood function $p(Y_t|\Theta)$. The endogenous predictor selection and learning render the state-space non-linear, making the analytic calculation of the likelihood intractable. In place of a Kalman Filter, we adopt the Bootstrap particle filter as described in Herbst and Schorfheide (2015). The particle filter is a way to produce recursive approximations to the distribution of the latent state variables X_t. Our algorithm samples the posterior distribution through an adapted Random-Walk Metropolis-Hastings (RWMH) MCMC technique with the particle-filter-based estimate of the likelihood function. A detailed discussion of the MCMC-particle filter algorithm, the priors, and its implementation is in Appendix C.

2.4.2 Parameter estimates

Table 1 reports the means and the fifth and ninety-fifth percentiles of the marginal posterior distributions of the parameters. In the estimation, we fixed the parameters $\beta = 0.99, s_b = 0.30$, and $\sigma = 2.00$. The parameter estimates are mostly in line with previous estimates in the literature, with a few notable exceptions described below.

We briefly discuss the key parameters associated with endogenously (non-)Ricardian beliefs. The "intensity of choice" parameter, ω, has a mean estimate of 6.301. This estimate is close to the value of 5.04 estimated by Cornea-Madeira et al. (2017) from a benchmark New Keynesian model with heterogeneous expectations. The speed of adjustment in estimating the relative forecasting accuracies of the two models, γ_2, has a mean estimate of 0.091, implying that model selection averages past forecast errors with a geometric decay of 0.909. The interpretation of this value is that the memory size for model selection is approximately 11 quarters. The 5/95% intervals imply a range of memory from 9.7-12.7 quarters. The mean estimated gain for coef-

[15] We end the sample before the ZLB episode as incorporating an effective lower bound on interest rates is beyond the scope of the present paper. Below, we consider the model's performance in an out of sample forecasting exercise.

Table 1: Posterior distribution of parameters

		90% credible region	
Parameter	Mean	5%	95%
Structural parameters			
κ	0.392	0.351	0.433
α	0.584	0.538	0.630
Policy parameters			
ϕ_π	1.854	1.665	2.042
ϕ_y	0.119	0.105	0.132
ϕ_b	0.062	0.034	0.091
Exogenous shocks			
ρ_g	0.050	0.007	0.093
ρ_u	0.343	0.296	0.389
ρ_w	0.305	0.262	0.347
ρ_z	0.860	0.814	0.907
$100\sigma_g$	0.631	0.429	0.833
$100\sigma_u$	0.141	0.002	0.281
$100\sigma_w$	1.440	1.359	1.526
$100\sigma_z$	0.235	0.117	0.353
Learning parameters			
ω	6.301	4.428	8.174
γ_1	0.002	0.000	0.003
γ_2	0.091	0.079	0.103

ficient updating, γ_1, is 0.002. This value is on the lower end of previously reported estimates. For instance, Eusepi and Preston (2018) reports a value of 0.035, closer to the high end, but for a different learning model and without the forecast model selection. The relative values for γ_1, γ_2 imply that these estimates suggest fast-slow learning dynamics. Model selection takes place faster than coefficient estimation in response to past data. As we see later, these gain parameters imply recurrent, but persistent, belief changes.

The estimated Taylor-rule policy parameters are in line with previous estimates (e.g., Eusepi and Preston (2018), Justiniano et al. (2011)) and suggest an active monetary policy. While the estimated reaction coefficients are higher than these earlier papers, the coefficients are smaller than the active monetary/passive fiscal regime in Bianchi and Ilut (2017). The fiscal policy reaction coefficient is close to the mean value reported in Eusepi and Preston (2018). The estimated shock processes are also similar to existing estimates. One difference is that the estimated process for the aggregate demand shock is less volatile and persistent than typical estimates. However, we include serially correlated monetary policy shocks, whose innovations have the highest variance among the set of shocks considered. The estimated fiscal shocks are also the most persistent.

Two parameters whose estimates differ somewhat from many estimates in the literature involve the slope of the Phillips curve and the degree of price rigidity in the economy. For the latter, we estimate that prices update, on average, every 2.4 quarters more frequently than the typical range of 4-7 quarters. This estimate is in line with Klenow and Kryvtsov (2008). Finally, the slope of the Phillips curve is on the high side and considerably higher than in Eusepi and Preston (2018).

We can assess the empirical relevance of the non-Ricardian belief mechanism via the Bayes factor comparing our model to a version that fixes the fraction n at the temporary (non-)Ricardian equilibrium, i.e., $n_t = \{0, 1\}$ in all periods. When we make this comparison, we find that the difference in log marginal likelihoods is in favor of the model with endogenously (non-)Ricardian beliefs and whose difference totals 5.7180 when $n = 0$ and 7.026 when $n = 1$. The Bayes factor takes the product of the ratios of the marginal likelihoods and the prior probabilities for each of the two models. The empirical evidence favoring the model with non-Ricardian beliefs is substantial in that the data will prefer that model over the Ricardian or fully non-Ricardian versions for any prior ratio below 304.29 and 1,125.40, respectively. Thus, a researcher would have to be *a priori* more confident in a rational expectations (i.e., self-confirming expectations) model by a factor of over 300 to reject the model with non-Ricardian beliefs.[16]

3 (Non-)Ricardian Beliefs: mechanics

An important component to the empirical findings is that the learning dynamics have a non-Ricardian equilibrium as its limit point, but also feature occasional departures (escapes) to the Ricardian restricted perceptions equilibrium. This section establishes two theoretical facts, key to understanding this mechanism. First, Ricardian beliefs are fragile in the sense that non-Ricardian effects arise for any $n > 0$. Second, one can always find an intensity of choice parameter ω so that a non-Ricardian equilibrium exists.

3.1 Fragility of Ricardian beliefs

We begin by establishing that Ricardian beliefs arise as a knife-edge result.

When $n = 0$ – all agents forecast with model-b – then a (weak) Ricardian result arises with fiscal shocks having only a transitory impact on inflation. For any $n > 0$, on the other hand, beliefs are non-Ricardian and fiscal shocks have a persistent inflationary effect. For expositional clarity, we will simplify the model slightly. Assume the only shock is an *iid* surplus shock z_t, $\phi_y = 0$, and let $s_b \to 0$, i.e., the long-run debt-output ratio is small. Each assumption is inconsequential to the result:

Proposition 1 *(Weak) Ricardian Equivalence obtains if and only if $n = 0$, that is all agents forecast with the debt-model.*

Proof. Notice in this simplified setting that

$$\begin{aligned}
b_{t+1} &= \beta^{-1}\left[b_t - s_t\right] \\
i_t &= \phi_\pi \pi_t \\
\hat{E}_t p^*_{t+1} &= n\psi_p^s s_t + (1-n)\psi_p^b b_t \\
\hat{E}_t v_{t+1} &= n\psi_v^s s_t + (1-n)\psi_v^b b_t.
\end{aligned}$$

[16] Of course, this model comparison is just on those beliefs within this paper's economic environment. One might have a subjective prior for a different environment and need to compute the associated marginal likelihood.

It follows that in a temporary equilibrium

$$\begin{aligned}\pi_t &= \xi_p b_t + \xi_z z_t \\ v_t &= \xi_v b_t + \tilde{\xi}_z z_t.\end{aligned}$$

The expressions for ξ are complicated functions of n and the ψ's, but simplify below.

For the moment, hold n fixed and focus on *extrinsic heterogeneity*. Within a restricted perceptions equilibrium,

$$\begin{aligned}Es_{t-1}\left\{\begin{bmatrix}\frac{\pi_t}{1-\alpha} \\ v_t\end{bmatrix} - \begin{bmatrix}\psi_p^s \\ \psi_v^s\end{bmatrix} s_{t-1}\right\} &= 0 \\ Eb_{t-1}\left\{\begin{bmatrix}\frac{\pi_t}{1-\alpha} \\ v_t\end{bmatrix} - \begin{bmatrix}\psi_p^b \\ \psi_v^b\end{bmatrix} b_{t-1}\right\} &= 0.\end{aligned}$$

Solving the orthogonality conditions leads to the belief coefficients

$$\begin{aligned}\psi_p^b &= \beta^{-1}\xi_p(1-\phi_b)/\left(1-\alpha\right) \\ \psi_v^b &= \beta^{-1}\xi_v(1-\phi_b),\end{aligned}$$

and,

$$\begin{aligned}\psi_p^s &= \frac{\phi_b \psi_p^b}{\beta^2 + 2\phi_b - 1} \\ \psi_v^s &= \frac{\phi_b \psi_v^b}{\beta^2 + 2\phi_b - 1}.\end{aligned}$$

Substituting in for ξ_p, ξ_v and solving, leads to restricted perceptions equilibrium laws of motion

$$\begin{aligned}\pi_t &= \xi_p\left(n\right) b_t + \xi_z\left(n\right) z_t \\ v_t &= \xi_v\left(n\right) b_t + \tilde{\xi}_z\left(n\right) z_t.\end{aligned}$$

We are now in a position to establish the claim. Let $n = 0$. Then,

$$\begin{aligned}\psi_p^b &= 0 \\ \psi_v^b &= -\beta^{-1}\left(1-\beta\right)\left(1-\phi_b\right).\end{aligned}$$

It follows that $\xi_p(0) = 0$ and

$$\pi_t = -\frac{(1-\beta)}{\beta\,(1+\kappa\sigma\phi_\pi)} z_t.$$

Fiscal shocks have a purely transitory impact on inflation. Furthermore,

$$\hat{E}_t v_{t+1} = -\beta^{-1}\,(1-\beta)\,(1-\phi_b) b_t.$$

The expected continuation value equals the expected present-value of future taxes, given by the current debt-level. Thus, when $n = 0$, higher stocks of government debt have no real impact and the output gap is

$$y_t = -\frac{(1-\beta)\,[1+\kappa\sigma\,(\phi_\pi - 1)]}{\beta\,(1+\kappa\sigma\phi_\pi)} z_t.$$

Now suppose that $n \to 0$. It turns out that $\xi_p(0), \xi_p'(0) > 0$. By continuity, $\xi_p(n) > 0$ and $\xi_p'(n) > 0$. Inflation now depends directly on the stock of debt, the expected continuation value deviates from the expected present-value of future taxes, and Ricardian Equivalence fails. ■

Ricardian beliefs are fragile because only when beliefs are exactly correct will households correctly anticipate the path of future surpluses. Otherwise, they treat the stock of government debt as real wealth, which impacts both inflation and the output gap. This failure of Ricardian beliefs occurs even for a vanishing fraction of agents forecasting with the surplus model. Outside of the $n = 0$ restricted perceptions equilibrium, both types of agents have non-Ricardian beliefs.

3.2 Regimes

Figure 1 illustrates the full breadth of regimes consistent with the model. Most of the literature focuses on a perfect knowledge/rational expectations setting. The assumed policy setting implies that perfect knowledge produces a Ricardian regime. This paper focuses, though, on restricted perceptions. When $n = 0$, the model features a Ricardian regime. The regime is weakly Ricardian, though, because fiscal shocks can still have a transitory impact. The Ricardian regime, though, is fragile as any heterogeneity in beliefs, i.e. $n > 0$, leads to a non-Ricardian regime.

Determining n endogenously determines which regime we might observe in equilibrium. A complete characterization of misspecification equilibria is beyond the scope of

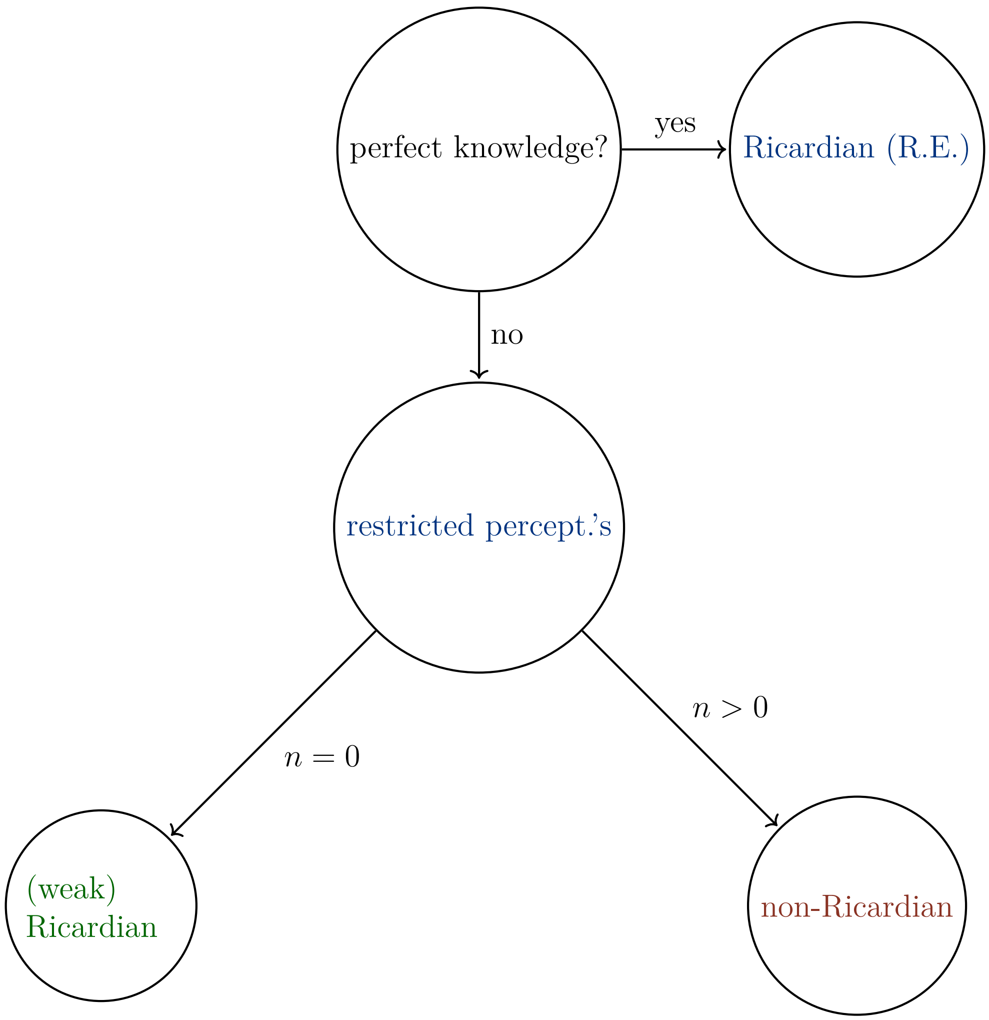


Figure 1: Classification of regimes.

this paper. However, it is possible to provide a sufficient condition for existence of the non-Ricardian regime.

Proposition 2 *For a sufficiently small intensity of choice ω, there exists a non-Ricardian misspecification equilibrium* $0 < \hat{n} < 1$.

Recall that the population distribution across models is given by the T-map $T_\omega(n)$, which is a function of the relative forecast accuracy of the surplus-based model. With the two forecast models producing forecast errors similar in magnitude (i.e. bounded) then, because of the continuity of T_ω, it is always possible to find an ω such that $0 < n < 1$. Essentially, with modest differences in forecasts between the two models we should expect agents to be distributed across both models. It turns out that this is the empirically relevant case. Our estimates imply an equilibrium value of $\hat{n} \approx 0.48$.

3.3 The Ricardian wedge

When $n = 0$ – everyone forecasts with model-b – beliefs are Ricardian. For any $n > 0$, beliefs are non-Ricardian. When agents are learning there are two sources for deviating from Ricardian beliefs: the choice of model, summarized by n_t, and coefficient updating $\psi^k_{j,t}, j = \{v, p\}, k = \{s, b\}$. Estimated parameters in Table 1 show that these components evolve at different rates γ_1, γ_2.

Thus, we define a function $\mathcal{R}\left(n, \left\{\psi^k_j\right\}, X_t\right)$ that provides a distance measure between aggregate beliefs and Ricardian beliefs. We define the *Ricardian wedge* by the function

$$\mathcal{R}\left(n, \left\{\psi^k_j\right\}_{j,k}, X_t\right) = \left(\hat{E}_t p^*_{t+1} - E^r_t p^*_{t+1}\right)^2 + \left(\hat{E}_t v_{t+1} - E^r_t v_{t+1}\right)^2,$$

where E^r denotes the expectations that would arise in a $n = 0$ world.[17] The Ricardian wedge is an appropriate distance measure for the impact of non-Ricardian beliefs on expectations about all payoff relevant aggregate variables.

4 (Non-)Ricardian beliefs: estimates

We turn to our main interest: measuring the impact of endogenously (non-)Ricardian beliefs on U.S. inflation. We begin with the estimated path for inflation and

[17]Below, it will be apparent that this wedge is isomorphic to one that calculates the distance to what a zero-mass fully rational agent would expect given the actual temporary equilibrium learning path.

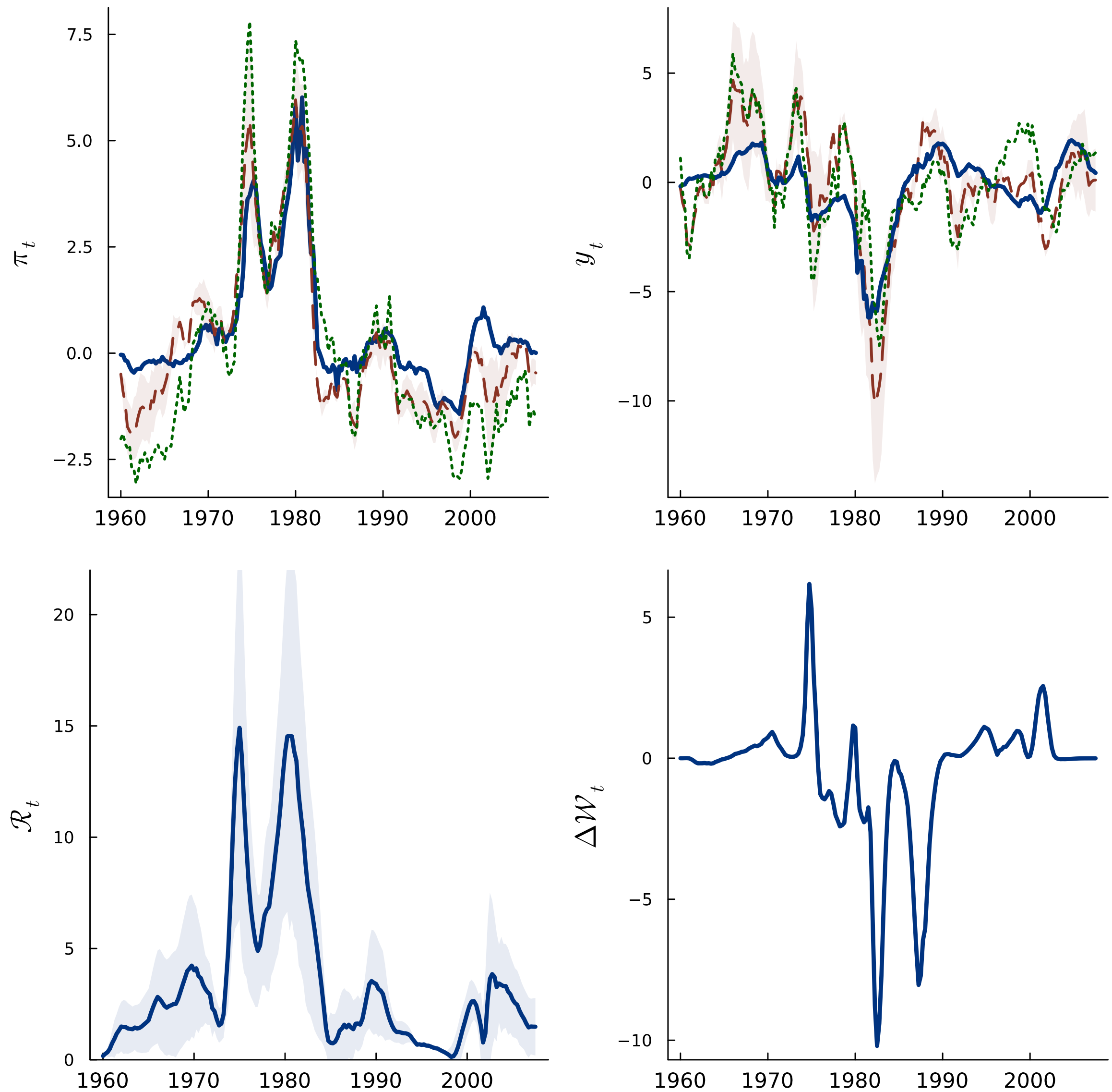
the output gap along with a decomposition into non-Ricardian effects. The evidence clearly points towards endogenous fluctuations between Ricardian and non-Ricardian regimes. We then study what drives these fluctuations by identifying the “most likely unlikely” sequence of shocks that can trigger an escape from a non-Ricardian to a Ricardian equilibrium. Counterfactual experiments deepen our understanding of the interaction between macroeconomic outcomes and beliefs.

4.1 Endogenously (non-)Ricardian beliefs in the U.S.

In this section, we provide compelling empirical evidence to support the main thesis of our paper: non-Ricardian beliefs play a significant role in explaining inflation dynamics, and the endogenous transition between non-Ricardian and Ricardian regimes can account for both periods of disinflation and low, stable inflation. We begin by estimating the (one step ahead) predicted state path $E\left(X_{t+1}|\mathbb{W}_t, \Theta\right)$ via the particle filter. From that predicted path, we can compute a decomposition of each data series into (non-)Ricardian effects as well as compute the Ricardian wedge. Our findings are presented in Figure 2 comprising four separate plots, each highlighting different aspects of U.S. macroeconomic data from 1960 to 2007. In the first two panels, dotted lines represent U.S. data, dashed lines denote model-predicted values, and solid lines measure the non-Ricardian effect by differencing model predicted values from a counterfactual with Ricardian beliefs. The shaded regions in each panel form the 60% confidence intervals, providing a clear visualization of the uncertainty around the estimates.

The first panel of the figure displays actual U.S. inflation rates, the model’s predicted inflation rate (computed as the one-step-ahead forecast from the particle filter), and a series that aids in decomposing the contribution to inflation from non-Ricardian effects. The non-Ricardian effect is measured by taking the difference between the predicted inflation rate and a counterfactual where the private-sector has full information rational expectations. This panel provides strong evidence that non-Ricardian effects significantly influenced inflation during the 1970s, with high inflation in the early 1970s being mostly non-Ricardian, and the high inflation at the end of the 1970s being estimated to be entirely non-Ricardian. In contrast, inflation during the 1960s was essentially Ricardian. The early 1980s disinflation was mostly Ricardian, while the late 1980s, a period of large budget deficits, saw modest non-Ricardian inflation. The 1990s, characterized by substantial budget surpluses, exhibited weak non-Ricardian effects that contributed to low inflation.

Figure 2: Estimated state dynamics. Dotted lines are U.S. data. Dashed lines are model-predicted values. Solid lines measure the non-Ricardian effect by differencing model-predicted values from a counterfactual with Ricardian beliefs. Shaded areas form the 60% credible interval.



Overall, the model's predictions capture the salient features of U.S. inflation rates. The model does not completely capture the volatility in inflation rates, particularly between 1960-1973. However, the non-Ricardian effect series also shows that learning and endogenously non-Ricardian beliefs captures substantially more of the variation in inflation than would occur, with the same sequence of shocks, under rational expectations.

The second panel presents the output gap, the model's predicted output gap, and the counterfactual between the model's prediction and rational expectations. It demonstrates that non-Ricardian effects on the output gap are more modest than those on inflation. During the 1960s and 1970s, non-Ricardian predicted output gaps did not correlate strongly with actual U.S. output gaps. In the late 1980s, the decomposition indicates that most of the predicted output gap resulted from non-Ricardian effects. However, during the 1990s, the model's output gap does not adequately explain U.S. output gaps. The early 2000s saw a strong non-Ricardian component in the output gap.

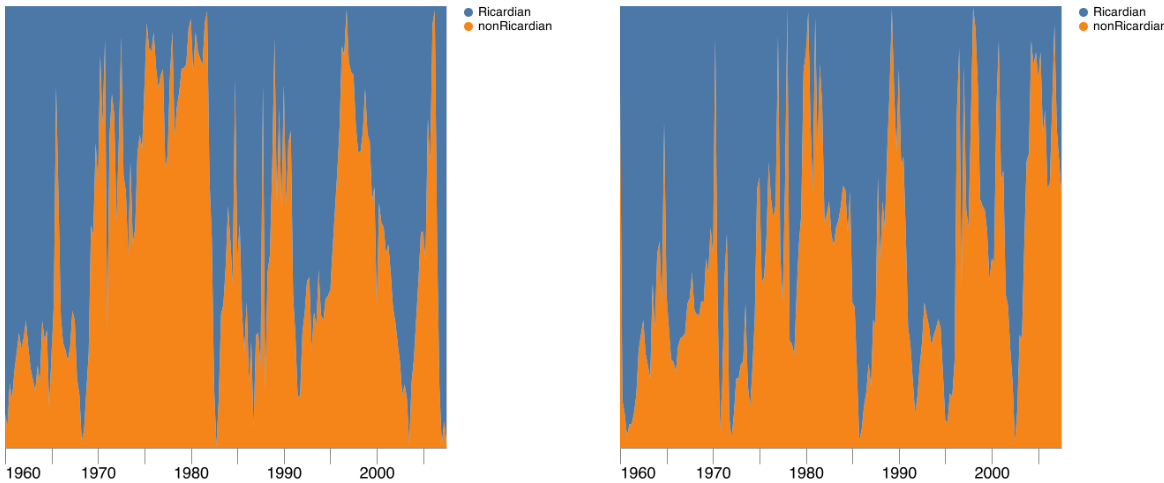
The third panel, displaying the estimated Ricardian wedges, corroborates these findings. The 1960s exhibit a wedge very close to zero, while a large wedge emerges from the 1970s until the early 1980s. Another significant wedge appears in the late 1980s, followed by a modest wedge after the 2001 recession. When the Ricardian belief wedge is large, predicted non-Ricardian components are also large.

The fourth panel, presenting the *perceived* non-Ricardian wealth gap, aligns with our findings on non-Ricardian effects and inflation. The perceived wealth gap comes from the aggregate Ricardian belief condition (2):

$$\mathcal{W}_t = b_t - \sum_{T \geq t} \beta^{T-t} \hat{E}_t \left\{ s_b \left[\beta i_T - \pi_T \right] - s_T \right\}.$$

The annuitized value ${}^{\mathcal{W}_t}/_{(1-\beta)}$ contributes directly to aggregate consumption. Under Ricardian beliefs, the perceived wealth gap is equal to zero. So the expression provides a direct measure of deviations in the perceived fiscal backing of debt (i.e., the present value of future surpluses) and the wealth effect arising from non-Ricardian beliefs.

For the 1960s and early 1970s, the perceived wealth wedge is estimated to be zero. At the end of the 1970s, households perceive a large positive wealth wedge, indicating that people interpreted high real government debt levels as real wealth and expected a correlation between inflation and debt. However, a strongly negative wealth



(a) Inflation. (b) Output gap.

Figure 3: Decomposing (non-)Ricardian effects on inflation and output gap. Height of the non-Ricardian bars represent the fraction attributable to non-Ricardian effects.

wedge occurs during the early 1980s disinflation, suggesting a correction to people's expectations about the present-value of future financial wealth. A similar negative wealth wedge occurs with the more modest disinflation in the early 1990s.

Figure 3 helps to visualize the non-Ricardian decomposition in Figure 2. In each panel, the box decomposes the fraction of the data series, in a given time period, explained by the estimated non-Ricardian effect. Figure 3a shows that the 1970s are strongly influenced by non-Ricardian effects. The non-Ricardian effects are a very small fraction of overall predicted inflation during the subsequent disinflation. Non-Ricardian effects on inflation appear again in the late 1980s and 1990s. Figure 3b similarly estimates regimes with (non-)Ricardian effects. These effects are not as strong or persistent as for inflation.

The most compelling part of these figures is that fluctuations between Ricardian and non-Ricardian regimes occur endogenously, in response to economic shocks, and without policy change. The next section turns to what shocks drive beliefs to fluctuate between belief regimes.

By closely examining the figure, readers can clearly see how the non-Ricardian effects have a significant impact on inflation dynamics, while endogenous transitions between non-Ricardian and Ricardian regimes help explain periods of disinflation and low, stable inflation. This strong evidence in favor of our thesis is further reinforced by the consistency between the estimated Ricardian wedges and the non-Ricardian wealth gap.

4.2 What drives Ricardian beliefs?

The model estimates indicate an equilibrium with non-Ricardian beliefs and $\hat{n} \approx 0.48$. This equilibrium is a limit point to the learning process. In finite time, though, learning dynamics can dislodge the economy from its stable equilibrium towards another point. The estimated paths in Figure 2 show that the economy periodically escapes its non-Ricardian equilibrium and a period of Ricardian beliefs emerge. According to the model estimates, non-Ricardian beliefs are the usual outcome, so the natural question is what drives beliefs to be (temporarily) Ricardian?

This question is similar in spirit to the one posed by Sargent (1999) and Cho et al. (2002). In their framework, there is no trade-off between inflation and unemployment. The central bank, however, entertains the possibility of a Phillips curve and chooses policy optimally given their estimates of the slope in their statistical Phillips curve. The high inflation time-consistent Kydland-Prescott outcome is a restricted perceptions equilibrium and stable under central bank learning. However, the right sequence of shocks can lead the central bank's estimates to the low inflation optimal policy. Eventually, time consistency re-emerges. Cho et al. (2002) and Williams (2019) develop tools to identify the "most likely unlikely sequence of shocks" that drives learning to escape the basin of attraction of the stable equilibrium.

We can apply a similar approach here to give insight into what types of shocks drive the non-Ricardian beliefs to become Ricardian, particularly for the period following the 1970s inflation. Following Cho et al. (2002) and Branch and Evans (2011), we replace the exogenous shocks with trinomial approximations. In particular, we assume that

$$\begin{aligned} g_t &\in \{-\sigma_g, 0, \sigma_g\} \\ u_t &\in \{-\sigma_u, 0, \sigma_u\} \\ w_t &\in \{-\sigma_w, 0, \sigma_w\} \\ z_t &\in \{-\sigma_z, 0, \sigma_z\} \end{aligned}$$

Then, we draw deterministic sequences of shocks from this set; these are the "unlikely sequences." For each sequence, initialize the model in equilibrium, and simulate the model for $T = 200$ periods. For each path, we calculate the non-Ricardian belief wedge. An escape (from the non-Ricardian equilibrium) occurs when either the p^*- component

Table 2: Escape paths to Ricardian beliefs.

escape time	shocks
13	$(\sigma_g, \sigma_u, 0, \sigma_z)$
21	$(\sigma_g, \sigma_u, \sigma_w, -\sigma_z)$
22	$(0, \sigma_u, \sigma_w, \sigma_z)$
22	$(0, -\sigma_u, -\sigma_w, -\sigma_z)$
22	$(-\sigma_g, -\sigma_u, -\sigma_w, \sigma_z)$
26	$(0, 0, \sigma_w, \sigma_z)$
26	$(\sigma_g, 0, \sigma_w, 0)$
26	$(-\sigma_g, 0, -\sigma_w, 0)$
28	$(0, -\sigma_u, -\sigma_w, 0)$
43	$(\sigma_g, -\sigma_u, \sigma_w, -\sigma_z)$

or the v- component of the wedge is below 10^{-5}. The escape time is the number of periods, for a given shock sequence, until an escape occurs.

Table 2 reports the escape paths. The first row, with the shortest escape time, is the dominant escape path. That is, the most likely unlikely sequence of shocks to drive the economy from a non-Ricardian to Ricardian equilibrium are positive demand and cost push shocks along with positive shocks to the budget surplus. We can interpret this to say that the escape from the 1970's non-Ricardian regime to the 1980's Ricardian regime was most likely the result of a sequence of shocks that reduces real government debt. While positive surplus shocks have an immediate effect, demand and price shocks increase inflation, which in turn reduces real government debt. Other shock sequences produce more circuitous paths. In each case, the Ricardian equilibrium is unstable under learning and so the paths under the repeated shock sequences eventually converge to a "pseudo" steady-state.

The intuition for the dominant escape path is clear. This transition to a Ricardian equilibrium occurs when individuals perceive that inflation and real government debt are uncorrelated. The sequence of dominant shocks results in drifting inflation coupled with a marked decrease in debt, primarily driven by repeated surplus shocks. However, under the dynamics of fast-slow learning, the adaptation of beliefs occurs gradually. These shocks lead agents using model-b to perceive a weaker correlation between inflation and debt. This altered perception, in turn, weakens the feedback loop from debt to inflation, reducing the regression coefficient of inflation on debt. Over time, this adjustment leads to model-b demonstrating a lower mean squared error. Given the faster model selection, this creates a self-confirming path to the Ricardian equilibrium. See

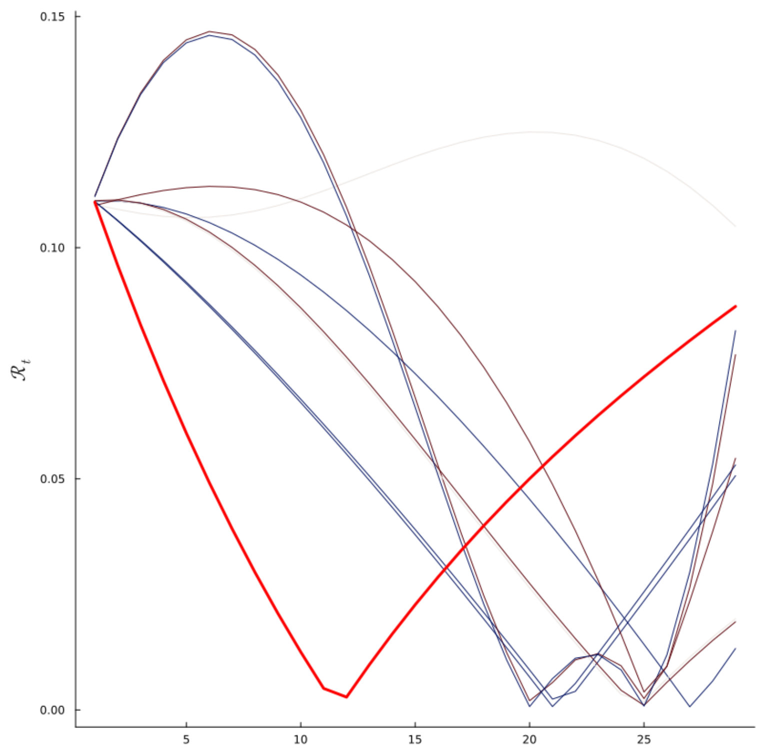


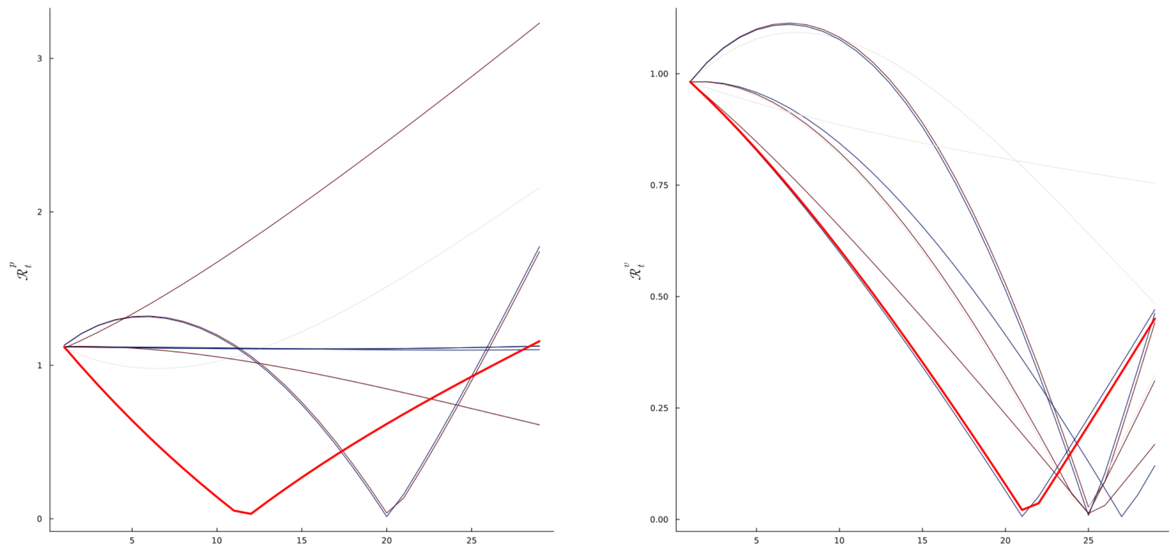
Figure 4: Escape dynamics to Ricardian equilibrium. Thick line denotes the dominant escape path (Table 2).

Figure 4 for all of the escape paths from the sequences in Table 2. Figure 5 decomposes the Ricardian wedges in Figure 4 into their price and continuation value components.

Examining the relationship between the Ricardian wedge and estimated shock sequences shows the dominant escape path at work in the estimated model. Figure 6 plots the predicted paths for the four shock sequences over the sample. The right-hand vertical axis measures the Ricardian wedge, overlayed in the panels as the dashed line.

Evidently, Ricardian wedges are positively correlated with the shocks g_t, u_t and w_t. Also, with the exception of the huge primary deficit in the mid 1970's, the cross-correlation between the wedge and z_t appears positive. Look at the period during the late 1970's to very early 1980's, a period estimated to be non-Ricardian. Here there is a sustained period of higher than average demand, supply and fiscal shocks. Table 2 and Figure 4 predict a persistent period of this shock pattern – the "most likely unlikely" sequence of shocks – will trigger an escape to the Ricardian equilibrium. That is, in fact, exactly what happens.

Again, the story that emerges is intuitive. Persistent positive shocks to demand,



(a) Inflation. (b) Continuation value.

Figure 5: Decomposing components of the Ricardian wedge along escape paths to the Ricardian equilibrium. Thick lines denote the dominant escape path.

mark-ups, and interest rates simultaneously with contractionary fiscal shocks generate negative correlation between inflation and government debt. This effect is mediated by expectations. An active monetary policy rule, though, predicts a weakening in that correlation. The inertia in beliefs delays the onset of the resulting escape to a Ricardian equilibrium after a sufficiently persistent sequence of shocks. As the effects of the unusual sequence of shocks fade, the non-Ricardian regime re-emerges.

4.3 Out-of-sample predictions

In this section, we delve into an out-of-sample exercise that offers valuable insights into our model's performance. We base our estimates on the 1960-2007 period, intentionally excluding the zero lower bound period from 2007.4 to 2015. Incorporating the zero lower bound would introduce unnecessary complexity to our already computationally challenging model and extend beyond the scope of this study. While we could employ a shadow fed funds rate to circumvent the zero lower bound, we find it more prudent to focus on out-of-sample predictions as an effective means of evaluating our model.

The out-of-sample exercise is conducted as follows: we first establish the structural parameters by fixing them at their median posterior value, using estimates derived from

Figure 6: Estimated shocks and escape paths to Ricardian beliefs. Solid line are shocks, dashed line is the non-Ricardian wedge.

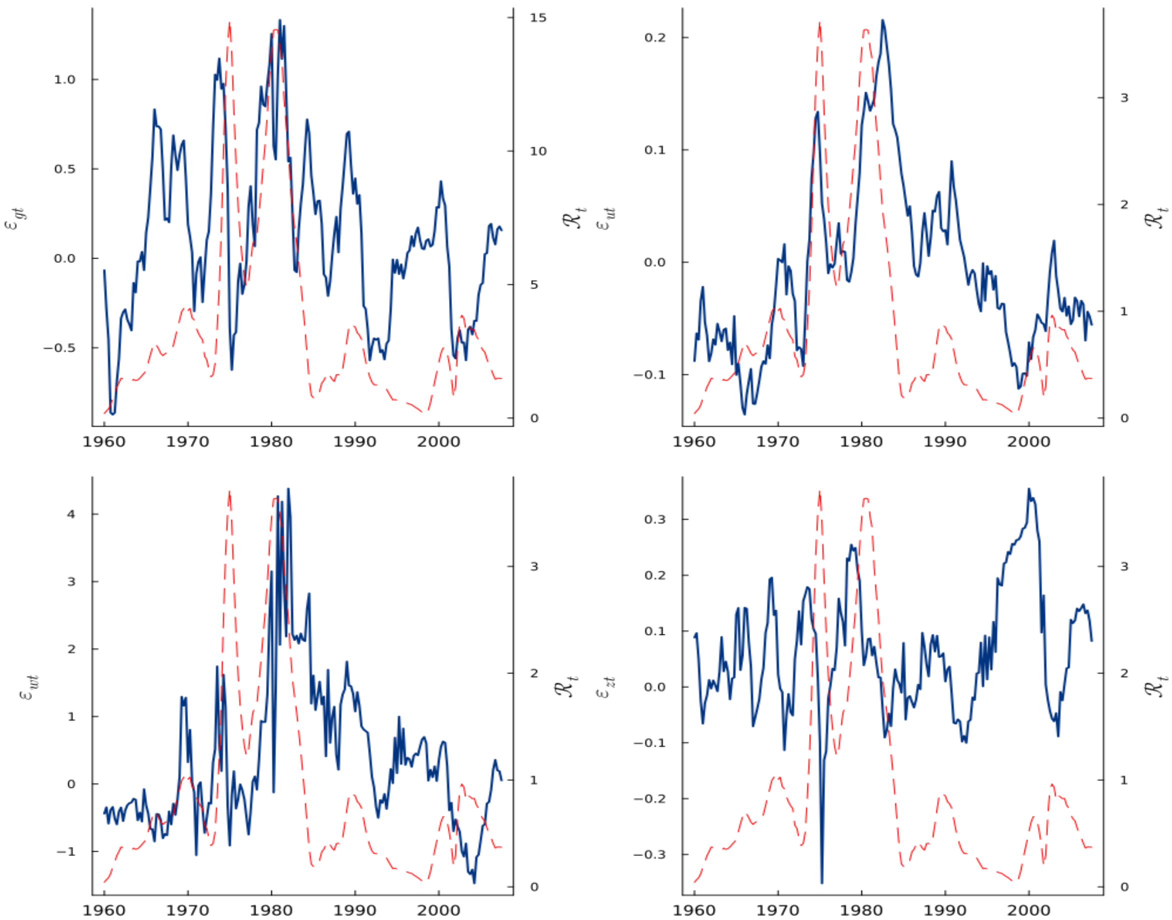
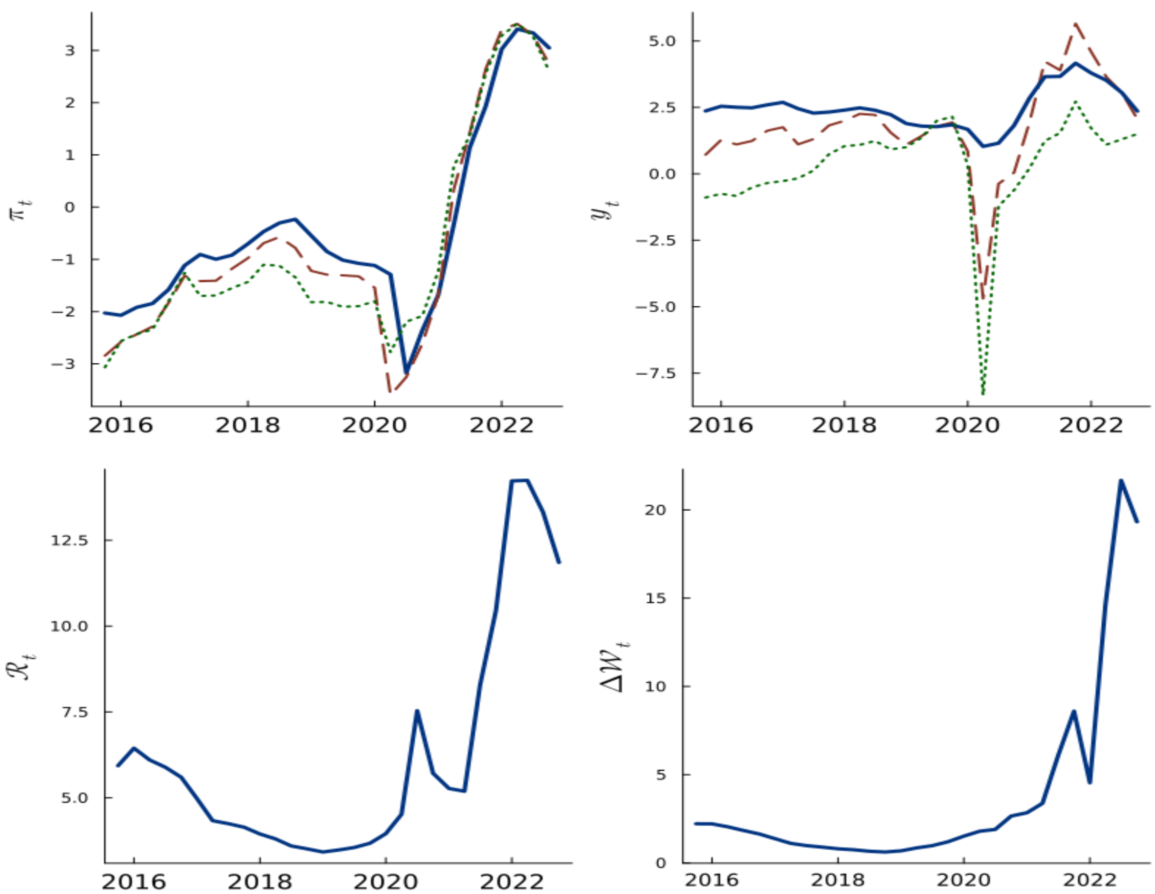


Figure 7: Out-of-sample model predictions 2016-2022. Dotted lines are U.S. data. Dashed lines are model-predicted values. Solid lines measure the non-Ricardian effect by differencing model-predicted values from a counterfactual with Ricardian beliefs.



the 1960-2015 sample. Next, we utilize the particle filter to generate one-period-ahead out-of-sample forecasts for inflation, the output gap, and the Ricardian wedge.

Figure 7 showcases the results of our out-of-sample exercise. The top two panels depict out-of-sample predictions for inflation and the output gap, offering a clear representation of our model's accuracy. Meanwhile, the bottom panel provides a glimpse into the predicted Ricardian wedge. Overall, our model demonstrates a strong alignment with the qualitative features of inflation and the output gap throughout this period. It is worth noting, however, that the model overestimates inflation between 2018 and 2020 and persistently overstates the output gap.

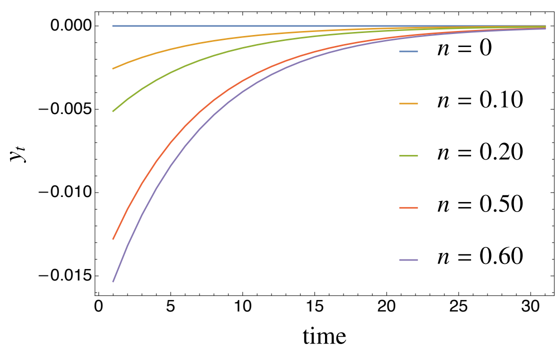
The bottom panel reveals an interesting aspect of our model's predictions: the period between 2017 and 2020 adheres to Ricardian beliefs. A sharp increase in the Ricardian wedge emerges in 2021, signifying a strong shift towards non-Ricardian beliefs. Thus, our analysis of the inflation and output gap plots indicates that the deflation experienced during the onset of the COVID pandemic was not driven by non-Ricardian beliefs.

Nevertheless, our out-of-sample forecasts identify a compelling development: starting in 2021, the significant increase in inflation over the 2021-2022 period appears to be largely attributable to endogenously non-Ricardian beliefs. The timing of this surge in non-Ricardian beliefs corresponds with the implementation of the American Rescue Plan's fiscal policies. This observation underscores the influential role non-Ricardian beliefs play in shaping macroeconomic outcomes and highlights the critical need to comprehend their interactions with fiscal and monetary policies.

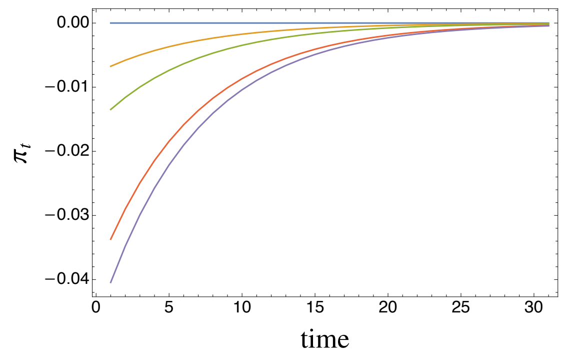
4.4 Counterfactuals

4.4.1 Extent of non-Ricardian beliefs

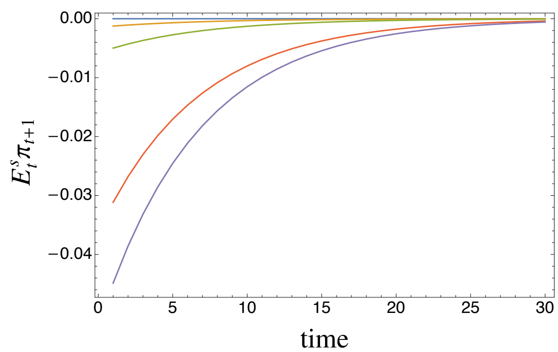
We can also gain more significant insights into the estimated mechanism by examining impulse responses to a contractionary fiscal policy shock. The impulse response function is non-linear, depending on the fraction of agents forecasting with the non-Ricardian model n and the recursively updated model coefficients and sequence of shocks. To present these impulse responses, we proceed by fixing n at the values within its estimated range for the median ($n \in [0, 0.60]$) and fixing the resulting coefficients at their RPE values. Figure 8 plots the results. Each (color-coded) line represents a different value for n, with the blue line representing $n = 0$, or the self-confirming fully Ricardian equilibrium.



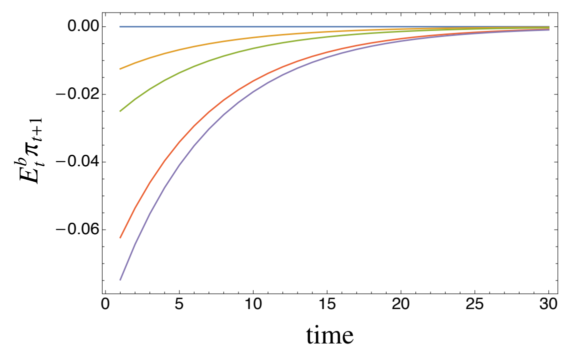
(a) output gap.



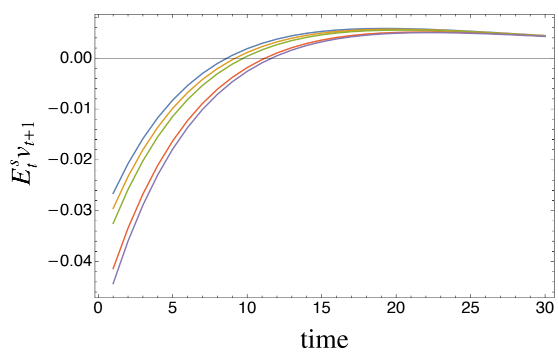
(b) inflation.



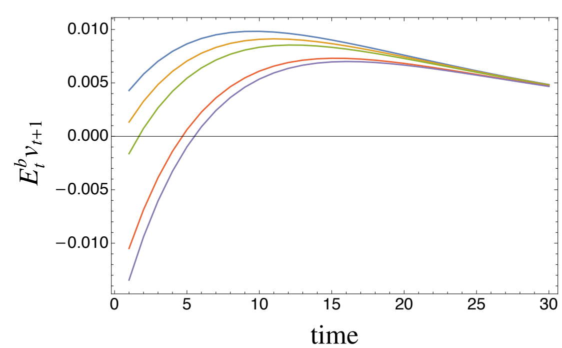
(c) expected inflation $E_t^s \pi_{t+1}$.



(d) expected inflation $E_t^b \pi_{t+1}$.



(e) expected continuation value $E_t^s v_{t+1}$.



(f) expected continuation value $E_t^b v_{t+1}$.

Figure 8: Impulse responses to a 1-unit contractionary fiscal shock. Each impulse response is computed for a different fraction of Ricardian beliefs, n, holding beliefs fixed at their RPE values.

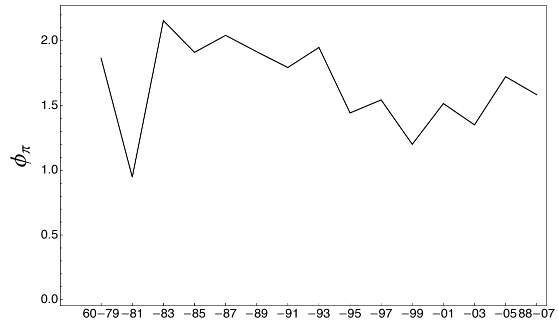
Evidently, from the top two panels, (a) and (b), the effect of an unanticipated contractionary fiscal shock depends on the degree of non-Ricardian beliefs at the time. When $n = 0$, the fiscal shocks have no impact on the output gap or inflation. Similarly, when more than half of the agents have non-Ricardian beliefs ($n \geq 0.5$), there is a more substantial economic reaction to the fiscal shock, with inflation and the output gap dropping. One can understand how expectations impact this finding by looking at the bottom two rows of Figure 8. The middle rows, (c) and (d), plot the inflation expectations by agents with the non-Ricardian and Ricardian forecast models, respectively. The bottom row plots these exact cross-type expectations for the continuation value of financial and non-financial wealth v_t. When the model is in a Ricardian equilibrium ($n = 0$), there is no impact on either agent type's expectations of future inflation. As seen in panel (b), these beliefs are self-confirming. However, the bottom panels show that the fiscal shock impacts expectations about future wealth as the real primary surplus changes expectations about future after-tax income. For agents with fully Ricardian beliefs (panel (f)), the unanticipated increase in the primary surplus leads them to expect higher future disposable income. However, in economies that are non-Ricardian, the presence of the non-Ricardian agents leads to non-Ricardian effects through the lower output gap, which mediates those expectations for the agents in panel (f). For the non-Ricardian expectations type, the positive surplus shock leads them always to expect lower continuation wealth in the near term. For all agents, as the economy recovers and the agents understand that the increase in surplus leads to an increase in disposable income, their expected continuation values v increase before returning to steady-state.

These impulse responses partially explain how the model generated higher inflation during the 1970s and anchored expectations during the 1990s. During the 1970s, the primary surplus shifted towards a sustained period of deficits. An increasing fraction of agents forecasting based on those surpluses helped generate inflation. During the 1980s, there were more significant primary deficits, but the switch toward Ricardian beliefs, and the inertia in the forecasting rules, mitigated those effects from impacting inflation. Conversely, during the early 1990s, fiscal policy turned toward sustained primary surpluses, and a fraction forecasting with the surplus model helped to reduce and anchor inflation expectations.[18]

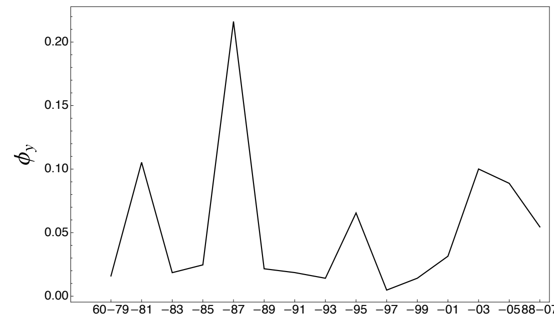
4.4.2 Policy change

The results thus far assume policy rules with time and state-invariant coefficients. However, a line of research argues that the stance of monetary and fiscal policy evolves over time. Davig and Leeper (2006), Bianchi (2013), and Bianchi and Ilut (2017) identify regime-switching policy rules that alternate between active/passive and pas-

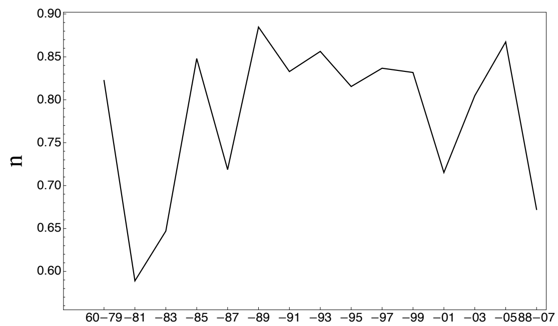
[18]Appendix 6 presents further evidence, external to the model, from the Survey of Professional Forecasters, that forecasts are consistent with a non-Ricardian regime in the 1990's.



(a) policy coefficient ϕ_π.

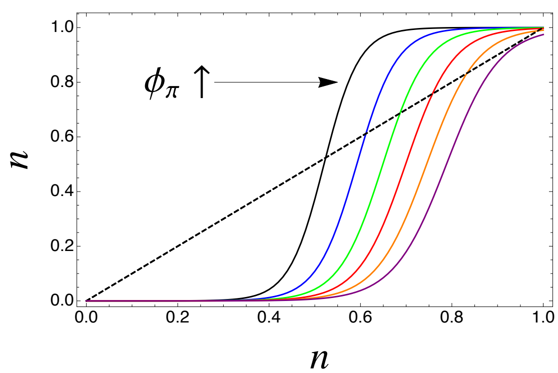


(b) policy coefficient ϕ_y.

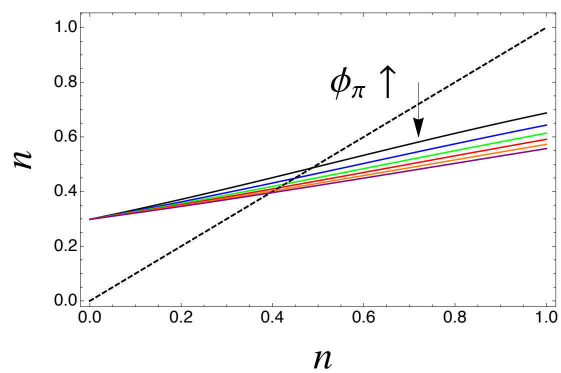


(c) fraction w/Ricardian beliefs.

(d) policy coefficient ϕ_b.



(e) $\omega = 100$.



(f) $\omega = 6.31$.

Figure 9: Rolling window estimation results. Panels (a)-(d) coefficient estimates. Panels (e)-(f) corresponding T-maps with $\omega = 100$ and $\omega = 6.03$, resp.

sive/active monetary/fiscal policies. Estimating regime-switching policy and beliefs is beyond this paper's scope, but we present rolling-window estimates of time-varying non-Ricardian beliefs alongside time-varying monetary and fiscal policies.

To make the analysis computationally feasible, we shut down endogenous selection of beliefs and estimate the fraction n as a structural parameter.[19] We consider rolling

[19] We can analytically evaluate the likelihood function with the Kalman Filter by shutting down the

20-year windows and use Bayesian techniques to estimate structural parameters, including n, over each rolling sample, advancing the window by eight quarters. This provides a rough estimate of policy and belief evolution.

Figure 9 shows the rolling window estimates for monetary policy rule reaction coefficients, which are consistent with previous findings. The period ending in 1981 featured a less active monetary policy, while subsequent windows were near two but declined during the late 1990s. The reaction coefficient to the output gap is estimated to be below 0.10 for most windows. Panel (d) plots the fiscal policy reaction coefficient ϕ_b, which increases through the mid-1980s before declining in later windows, corresponding to a relatively more active fiscal policy.

To better understand estimated beliefs in panel (c), we examine how changes in monetary policy rule (ϕ_π) or fiscal policy rule (ϕ_b) impact endogenous model selection in panels (e) and (f). Changes in the monetary policy rule have a modest impact on the equilibrium fraction of non-Ricardian beliefs in the estimated model version. More active monetary policy rules shrink the basin of attraction for non-Ricardian equilibria. Similar comparative statics arise for ϕ_y, but the effect from ϕ_b is non-monotonic.

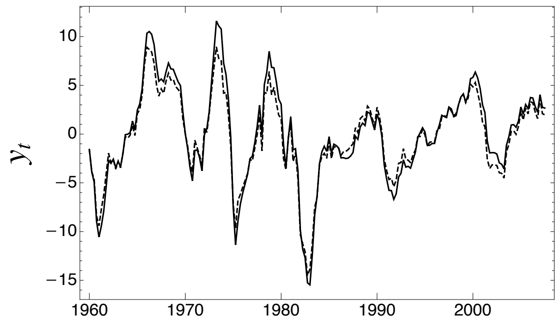
Returning to panel (c) of Figure 9, the average degree of non-Ricardian beliefs is significantly higher than in the benchmark estimation. Over the windows estimated with declining ϕ_π and ϕ_b, there is a downward trend in n. During the 1970s, policy coefficients moved ambiguously regarding their theoretical impact on n. Estimates are consistent with less active monetary policy at the beginning of the Great Inflation period, followed by an increasing role of non-Ricardian beliefs in maintaining high inflation.

In conclusion, we examine the policy implications of non-Ricardian beliefs and their association with economic instability. The answer depends on the policy stance leading to more non-Ricardian beliefs and its impact on the endogenous response of inflation and the output gap. Our analysis reveals that non-Ricardian beliefs are not inherently destabilizing, and policy coordination is possible even within an active/passive policy regime.

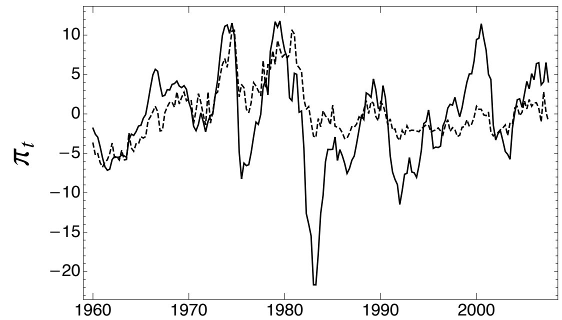
4.4.3 Learning speed

We use two other counterfactuals to disentangle the role played by real-time learning. The first counterfactual, plotted in Panels 10a and 10b, fixes the shocks and the parameters at their mean values except setting $\gamma_1 = 0.03$. This counterfactual undoes

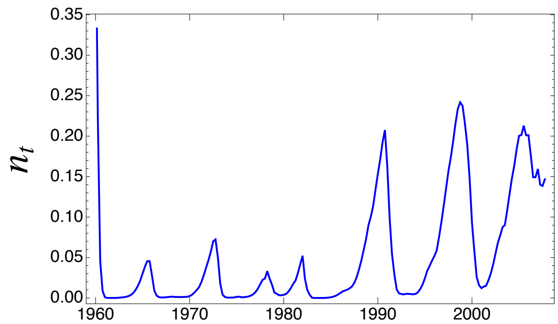
real-time predictor selection and focusing on the RPE in each window.



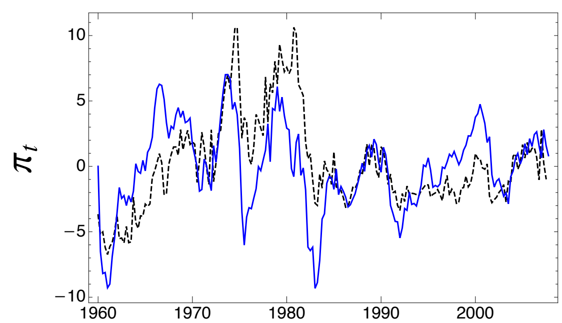
(a) fast learning ($\gamma_1 = 0.03$): output gap.



(b) fast learning ($\gamma_1 = 0.03$): inflation.



(c) no-learning: fraction non-Ricardian.



(d) no-learning: inflation.

Figure 10: Counterfactuals. Panels (a)-(b) from counterfactual with higher learning gain $\gamma_1 = 0.03$. Panels (c) and (d) from counterfactual with no learning. Solid lines are the counterfactual paths, dashed lines are the estimated paths.

the estimated fast-slow learning dynamic and considers the implications for the output gap and inflation if learning is fast. The right panel shows that inflation volatility is counterfactually high. Inflation is predicted in this alternate economy to be greater during the 1960s, with vast over-shooting during the 1970s: the counterfactual implication is a dramatic collapse from the 1970s inflation. The counterfactual scenario also predicts that faster learning would have led to a second high inflation episode during the late 1990s.

Panels 10c and 10d construct a counterfactual scenario that shuts down learning entirely. Parameters and shocks are held fixed at their estimated mean values, allowing n to evolve endogenously, but for each value of n_t, the model coefficients equal their RPE values. The left panel plots the counterfactual fraction with non-Ricardian beliefs, while the right panel is the corresponding inflation path. Clearly, in this scenario, beliefs tend to be more Ricardian but with more volatility in the later sample periods. The counterfactual scenario substantially under-predicts inflation during the 1970s and early 1980s and over-predicts during the 1990s. The model with no learning does well to explain inflation during the late 1980s and 2000s.

5 Related literature

This paper is related to a large literature that examines monetary policy design when replacing rational expectations with an adaptive learning rule. Key contributions include Bullard and Mitra (2002), Evans and Honkapohja (2003), and Preston (2005). The first to characterize fiscal and monetary policy interaction under adaptive learning is Evans and Honkapohja (2007) and Eusepi and Preston (2012). Evans et al. (2012) examine the conditions under which Ricardian equivalence holds or fails under adaptive learning. The theory of restricted perceptions is related to a wide variety of applications of misspecified models, e.g., Sargent (1999), Adam (2005), Branch and Evans (2006), Sargent (2008), Branch and McGough (2018), and Cho and Kasa (2015). Misperceptions and learning about policy rules are introduced by Cogley et al. (2015) and Hollmayr and Matthes (2015). This paper builds on insight from Woodford (2013), where an example of a restricted perceptions equilibrium leads to a failure of Ricardian equivalence. In short, our paper extends the theory of forecast misspecification in Branch and Evans (2006) into the Eusepi and Preston (2018) environment with fiscal and monetary policy interaction and generalizing the beliefs in Woodford (2013).

Our paper is also related to a long-standing tradition of constructing equilibria with the property that inflation is (partly) driven by fiscal policy. In the original contribution by Leeper (1991), an active fiscal policy, combined with a monetary policy not committed to price stability, generates inflation driven by fiscal variables, i.e., the "fiscal theory of the price-level." See also, Sims (1994), Cochrane (2001) and Woodford (2001). Recent research explains post-war U.S. inflation via recurrent change between non-Ricardian and Ricardian policy regimes. Examples include Davig and Leeper (2006), Sims (2011), and Bianchi and Ilut (2017). These papers also derive their results from an important role given to non-Ricardian beliefs. When agents assign a positive probability to changes from the Ricardian policy regime to the non-Ricardian policy regime, then the beliefs imply failure of Ricardian equivalence, and inflation is also a fiscal phenomenon. Non-Ricardian beliefs, though, in these settings require that agents anticipate recurrent changes in the coordination and objectives of fiscal and monetary policymakers. Indeed, there is not an empirical consensus for economically meaningful time-variation in policy rules (c.f. Sims and Zha, 2006; Primiceri, 2005; Liu et al., 2011). The question addressed by our paper is whether fiscal shocks can matter at times, even without policy regime change. Including the possibility of regime change is an interesting topic to be explored. We note, however, that our results do not suggest that policy regime change is an unimportant part of the inflation story. More subtle changes within the Ricardian policy regime can generate belief-driven regime change. We explored this point by taking rolling window estimates of the policy coefficients and non-Ricardian beliefs. We found that the estimated fraction of non-Ricardian beliefs evolves along with policy rules as predicted by theory. While beyond the scope of the present study, we leave open quantifying policy-regime change versus belief-regime change.

Finally, our theory here is inspired by and builds on Eusepi and Preston (2018), who show that replacing rational expectations with an adaptive learning rule produces temporary equilibrium dynamics that feature departures from Ricardian equivalence. In addition, their paper illustrates how the maturity structure of government debt has important implications for inflation in a non-Ricardian belief economy. They also estimate a quantitative version of their model and conduct counterfactual analyses demonstrating that perceived net wealth may be a significant factor in high-debt economies.

6 Conclusion

In this paper, we have studied the role of non-Ricardian beliefs in shaping macroeconomic outcomes, particularly inflation and output gap dynamics. We developed a model incorporating agents with Ricardian and non-Ricardian beliefs and provided a novel learning mechanism for agents to choose between these beliefs endogenously. Our approach sheds light on the interaction between agents' expectations, fiscal policy, and monetary policy, offering insights into the evolution of inflation and output gap dynamics over time.

Our main findings show that non-Ricardian beliefs played a significant role in generating higher inflation during the 1970s and anchoring expectations during the 1990s. We demonstrated that the presence of non-Ricardian beliefs can lead to more substantial economic reactions to fiscal shocks. Moreover, we found that the specific reaction coefficients in monetary and fiscal policies can vary over time and have an impact on the extent of non-Ricardian beliefs in the economy.

In addition, we investigated the policy implications of non-Ricardian beliefs and their association with economic instability. Our results revealed that non-Ricardian beliefs are not inherently destabilizing, and the impact of these beliefs on economic stability depends on the policy stance leading to their prevalence. This finding highlights the importance of policy coordination, even within an active/passive policy regime. A more active monetary policy stance increases the likelihood of a Ricardian regime.

This paper contributes to the literature by offering a richer understanding of the relationship between agents' expectations, fiscal policy, and monetary policy in shaping macroeconomic outcomes. Future research could extend the analysis by incorporating additional factors, such as incorporating heterogeneous agents with different types of non-Ricardian beliefs. For instance, Branch et al. (2022) show that restricted perceptions equilibria are often accompanied by similar equilibria that include a volatile sentiments, or sunspot, component. Another potential direction for future research is to explore the implications of our findings for optimal policy design, considering the interactions between monetary and fiscal policies and agents' endogenous belief formation.

References

Adam, K. (2005). Learning To Forecast And Cyclical Behavior Of Output And Inflation. *Macroeconomic Dynamics*, 9(1):1–27.

Andre, P., Pizzinelli, C., Roth, C., and Wohlfart, J. (2022). Subjective models of the macroeconomy: Evidence from experts and representative samples. *Review of Economic Studies*, 89(6):2958–2991.

Bassetto, M. (2002). A Game-Theoretic View of the Fiscal Theory of the Price Level. *Econometrica*, 70(6):2167–2195.

Bianchi, F. (2013). Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics. *Review of Economic Studies*, 80(2):463–490.

Bianchi, F. and Ilut, C. (2017). Monetary/Fiscal Policy Mix and Agents' Beliefs. *Review of Economic Dynamics*, 26:113–139.

Branch, W. A. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations. *Economic Journal*, 114(497):592–621.

Branch, W. A. and Evans, G. W. (2006). Intrinsic Heterogeneity in Expectation Formation. *Journal of Economic Theory*, 127(1):264–295.

Branch, W. A. and Evans, G. W. (2011). Learning about risk and return: A simple model of bubbles and crashes. *American Economic Journal: Macroeconomics*, 3(3):159–191.

Branch, W. A. and McGough, B. (2009). A New Keynesian Model with Heterogeneous

Expectations. *Journal of Economic Dynamics and Control*, 33(5):1036–1051.

Branch, W. A. and McGough, B. (2018). Heterogeneous Expectations and Micro-Foundations in Macroeconomics. In Hommes, C. H. and LeBaron, B., editors, *Handbook of Computational Economics*, volume 4. Elsevier, Amsterdam.

Branch, W. A., McGough, B., and Zhu, M. (2022). Statistical sunspots. *Theoretical Economics*, 17(1):291–329.

Brock, W. A. and Hommes, C. H. (1997). A Rational Route to Randomness. *Econometrica*, 65(5):1059–1095.

Bullard, J. B. and Mitra, K. (2002). Learning about Monetary Policy Rules. *Journal of Monetary Economics*, 49(6):1105–1129.

Calvo, G. A. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics*, 12(3):383–398.

Cho, I.-K. and Kasa, K. (2015). Learning and Model Validation. *Review of Economic Studies*, 82(1):45–82.

Cho, I.-K., Williams, N., and Sargent, T. J. (2002). Escaping Nash Inflation. *Review of Economic Studies*, 69:1–40.

Cochrane, J. H. (2001). Long-Term Debt and Optimal Policy in the Fiscal Theory of the Price Level. *Econometrica*, 69(1):69–116.

Cogley, T., Matthes, C., and Sbordone, A. (2015). Optimized taylor rules for disinflation when agents are learning. *Journal of Monetary Economics*, 72:131–147.

Cornea-Madeira, A., Hommes, C. H., and Massaro, D. (2017). Behavioral Heterogeneity in U.S. Inflation Dynamics. *Journal of Business and Economic Statistics*, forthcoming.

Davig, T. and Leeper, E. M. (2006). Fluctuating Macro Policies and the Fiscal Theory. In Acemoğlu, D., Rogoff, K. S., and Woodford, M., editors, *NBER Macroeconomics Annual*, volume 21, pages 247–316. MIT Press, Cambridge, MA.

Dupuis, P. and Ellis, R. S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York.

Eusepi, S. and Preston, B. (2012). Debt, Policy Uncertainty and Expectations Stabilization. *Journal of the European Economic Association*, 10(4):860–886.

Eusepi, S. and Preston, B. (2018). Fiscal Foundations of Inflation: Imperfect Knowledge. *American Economic Review*, 108(9):2551–2589.

Evans, G. W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, New Jersey.

Evans, G. W. and Honkapohja, S. (2003). Expectations and the Stability Problem for Optimal Monetary Policies. *Review of Economic Studies*, 70(4):807–824.

Evans, G. W. and Honkapohja, S. (2007). Policy Interaction, Learning, and the Fiscal Theory of Prices. *Macroeconomic Dynamics*, 11(5):665–690.

Evans, G. W., Honkapohja, S., and Mitra, K. (2012). Does Ricardian Equivalence Hold When Expectations are not Rational? *Journal of Money, Credit and Banking*, 44(7):1259–1283.

Evans, G. W., Honkapohja, S., and Williams, N. (2010). Generalized Stochastic Gradient Learning. *International Economic Review*, 51:237–262.

Favero, C. A. and Giavazzi, F. (2012). Measuring Tax Multipliers: the Narrative Method in Fiscal VARs. *American Economic Journal: Economic Policy*, 4(2):69–94.

Herbst, E. P. and Schorfheide, F. (2015). *Bayesian Estimation of DSGE Models*. Princeton University Press.

Hollmayr, J. and Matthes, C. (2015). Learning about fiscal policy and the effects of policy uncertainty. *Journal of Economic Dynamics and Control*, 59:142–162.

Justiniano, A., Primiceri, G. E., and Tambalotti, A. (2011). Investment Shocks and the Relative Price of Investment. *Review of Economic Dynamics*, 14(1):102–121.

Klenow, P. J. and Kryvtsov, O. (2008). State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation? *Quarterly Journal of Economics*, 123(3):863–904.

Leeper, E. M. (1991). Equilibria under Active and Passive Monetary and Fiscal Policies. *Journal of Monetary Economics*, 27(1):129–147.

Liu, Z., Waggoner, D. F., and Zha, T. (2011). Source of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach. *Quantitative Economics*, 2(2):251–301.

Marcet, A. and Sargent, T. J. (1989). Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models. *Journal of Economic Theory*, 48:337–368.

Milani, F. (2007). Expectations, Learning and Macroeconomic Persistence. *Journal of Monetary Economics*, 54(7):2065–2082.

Preston, B. (2005). Learning about Monetary Policy Rules when Long-Horizon Expectations Matter. *International Journal of Central Banking*, 1(2):81–126.

Primiceri, G. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies*, 72:821–852.

Sargent, T. J. (1999). *The Conquest of American Inflation*. Princeton University Press, Princeton, NJ.

Sargent, T. J. (2008). Evolution and Intelligent Design. *American Economic Review*, 98(1):5–37.

Sargent, T. J. and Williams, N. (2005). Impacts of Priors on Convergence and Escapes from Nash Inflation. *Review of Economic Dynamics*, 8:360–391.

Sargent, T. J., Williams, N., and Zha, T. (2006). Shocks and Government Beliefs: The Rise and Fall of American Inflation. *American Economic Review*, 96(4):1193–1224.

Sims, C. A. (1994). A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy. *Economic Theory*, 4(3):381–399.

Sims, C. A. (2011). Stepping on a Rake: the Role of Fiscal Policy in the Inflation of the 1970s. *European Economic Review*, 55(1):48–56.

Sims, C. A. and Zha, T. (2006). Were there Regime Switches in US Monetary Policy? *American Economic Review*, 96:54–81.

Taylor, J. B. (1993). Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214.

Williams, N. (2019). Escape Dynamics in Learning Models. *Review of Economic Studies*, 86(2):882–912.

Woodford, M. (2001). Fiscal Requirements for Price Stability. *Journal of Money, Credit and Banking*, 33(3):669–728.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press, Princeton, NJ.

Woodford, M. (2013). Macroeconomic Analysis Without the Rational Expectations Hypothesis. *Annual Review of Economics*, 5:303–346.

Young, P. H. (2004). *Strategic Learning and Its Limits.* Oxford University Press, Oxford, UK.

A Appendix

A.1 Computation of the restricted perceptions equilibrium

For a given distribution of PLMs, n, for all versions of the model the RPE can be computed in a similar way. First, we can re-organize the ALM to obtain

$$y_t = \delta_0 b_{t+1} + \delta_1 b_t + \delta_2 s_t + \delta_3 u_t \tag{A.1.1}$$

$$b_{t+1} = \xi_1 b_t + \xi_2 s_t + \xi_3 u_t. \tag{A.1.2}$$

Moreover, we can aggregate (6.3) and combine it with (5), (6.5), (A.1.1) and (A.1.2) to obtain

$$v_t = \mu_{v,1} b_t + \mu_{v,2} s_t + \mu_{v,3} u_t, \tag{A.1.3}$$

and (6.5), (4), (A.1.1) and (A.1.2) imply that

$$p_t^* = \mu_{p,1} b_t + \mu_{p,2} s_t + \mu_{p,3} u_t. \tag{A.1.4}$$

We combine (A.1.3) and (A.1.4) to

$$\mathbf{z}_t = \boldsymbol{\mu}_b b_t + \boldsymbol{\mu}_s s_t + \boldsymbol{\mu}_u u_t, \tag{A.1.5}$$

where $\mathbf{z}_t \equiv (v_t, p_t^*)'$, $\boldsymbol{\mu}_b \equiv (\mu_{v,1}, \mu_{p,1})'$, $\boldsymbol{\mu}_s \equiv (\mu_{v,2}, \mu_{p,2})'$, and $\boldsymbol{\mu}_u \equiv (\mu_{v,3}, \mu_{p,3})'$.

Next, recall that PLMs are given by

$$\mathbf{z}_t = \psi^s s_{t-1} + \eta_t \tag{A.1.6}$$

$$\mathbf{z}_t = \psi^b b_{t-1} + \eta_t, \tag{A.1.7}$$

where $\psi^s \equiv (\psi_v^s, \psi_p^s)'$, $\psi^b \equiv (\psi_v^b, \psi_p^b)'$ and $\eta_t \equiv (\eta_{v,t}, \eta_{p,t})'$. This implies four orthogonality conditions that can be written as

$$0 \overset{!}{=} E[s_{t-1}\eta_t] = E[s_t \eta_{t+1}] \tag{A.1.8}$$

$$0 \overset{!}{=} E[b_{t-1}\eta_t] = E[b_t \eta_{t+1}]. \tag{A.1.9}$$

Now, plug the PLM (A.1.6) and ALM (A.1.5) into (A.1.8), i.e.,

$$\begin{aligned}0 &\overset{!}{=} E[s_t\eta_{t+1}] = E[s_t(\mathbf{z}_{t+1} - \psi^s s_t)]\\ \Leftrightarrow \psi^s E[s_t^2] &= E[s_t\mathbf{z}_{t+1}].\end{aligned}$$

Equation by equation, we obtain

$$\begin{aligned}\Leftrightarrow \psi_v^s E[s_t^2] &= E\left[s_t\left(\mu_{v,1}b_{t+1} + \mu_{v,2}s_{t+1} + \mu_{v,3}u_{t+1}\right)\right]\\ \psi_v^s E[s_t^2] &= \mu_{v,1}E\left[s_t b_{t+1}\right] + \mu_{v,2}E\left[s_t s_{t+1}\right] + \mu_{v,3}E\left[s_t u_{t+1}\right]\\ \Leftrightarrow \psi_v^s &= \mu_{v,1}\frac{E[s_t b_{t+1}]}{E[s_t^2]} + \mu_{v,2}\frac{E[s_t s_{t+1}]}{E[s_t^2]} + \mu_{v,3}\frac{E[s_t u_{t+1}]}{E[s_t^2]} \qquad \text{and} \quad \text{(A.1.10)}\end{aligned}$$

$$\begin{aligned}\Leftrightarrow \psi_p^s E[s_t^2] &= E\left[s_t\left(\mu_{p,1}b_{t+1} + \mu_{p,2}s_{t+1} + \mu_{p,3}u_{t+1}\right)\right]\\ \psi_p^s E[s_t^2] &= \mu_{p,1}E\left[s_t b_{t+1}\right] + \mu_{p,2}E\left[s_t s_{t+1}\right] + \mu_{p,3}E\left[s_t u_{t+1}\right]\\ \Leftrightarrow \psi_p^s &= \mu_{p,1}\frac{E[s_t b_{t+1}]}{E[s_t^2]} + \mu_{p,2}\frac{E[s_t s_{t+1}]}{E[s_t^2]} + \mu_{p,3}\frac{E[s_t u_{t+1}]}{E[s_t^2]}. \qquad \text{(A.1.11)}\end{aligned}$$

Likewise plug the PLM (A.1.7) and ALM (A.1.5) into (A.1.9), i.e.,

$$\begin{aligned}0 &\overset{!}{=} E[b_t\eta_{t+1}] = E[b_t(\mathbf{z}_{t+1} - \psi^b b_t)]\\ \Leftrightarrow \psi^b E[b_t^2] &= E[b_t\mathbf{z}_{t+1}].\end{aligned}$$

Again, equation by equation, we obtain

$$\begin{aligned}\Leftrightarrow \psi_v^b E[b_t^2] &= E\left[b_t\left(\mu_{v,1}b_{t+1} + \mu_{v,2}s_{t+1} + \mu_{v,3}u_{t+1}\right)\right]\\ \psi_v^b E[b_t^2] &= \mu_{v,1}E\left[b_t b_{t+1}\right] + \mu_{v,2}E\left[b_t s_{t+1}\right] + \mu_{v,3}E\left[b_t u_{t+1}\right]\\ \Leftrightarrow \psi_v^b &= \mu_{v,1}\frac{E[b_t b_{t+1}]}{E[b_t^2]} + \mu_{v,2}\frac{E[b_t s_{t+1}]}{E[b_t^2]} + \mu_{v,3}\frac{E[b_t u_{t+1}]}{E[b_t^2]} \qquad \text{and} \quad \text{(A.1.12)}\\ \Leftrightarrow \psi_p^b E[b_t^2] &= E\left[b_t\left(\mu_{p,1}b_{t+1} + \mu_{p,2}s_{t+1} + \mu_{p,3}u_{t+1}\right)\right]\\ \psi_p^b E[b_t^2] &= \mu_{p,1}E\left[b_t b_{t+1}\right] + \mu_{p,2}E\left[b_t s_{t+1}\right] + \mu_{p,3}E\left[b_t u_{t+1}\right]\\ \Leftrightarrow \psi_p^b &= \mu_{p,1}\frac{E[b_t b_{t+1}]}{E[b_t^2]} + \mu_{p,2}\frac{E[b_t s_{t+1}]}{E[b_t^2]} + \mu_{p,3}\frac{E[b_t u_{t+1}]}{E[b_t^2]}. \qquad \text{(A.1.13)}\end{aligned}$$

The next step is to compute the moments. For this purpose, it is convenient to combine (A.1.2) and (6) in a VAR(1), i.e.,

$$\begin{bmatrix} 1 & -\xi_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_{t+1} \\ s_t \end{bmatrix} = \begin{bmatrix} \xi_1 & 0 \\ \phi_b & 0 \end{bmatrix} \begin{bmatrix} b_t \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \xi_3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ z_t \end{bmatrix}$$
$$\Leftrightarrow \mathcal{Y}_t = \mathbf{A}\mathcal{Y}_{t-1} + \mathbf{C}\varepsilon_t, \tag{A.1.14}$$

where $\mathcal{Y}_t \equiv (b_{t+1}, s_t)'$ and $\varepsilon_t \equiv (u_t, z_t)'$.

Define the variance-covariance matrix $\mathbf{\Omega} \equiv E[\mathcal{Y}_t\mathcal{Y}_t']$ and likewise $\mathbf{\Sigma} \equiv E[\varepsilon_t\varepsilon_t']$. Then we can compute

$$\begin{aligned} \mathbf{\Omega} &= E[(\mathbf{A}\mathcal{Y}_{t-1} + \mathbf{C}\varepsilon_t)(\mathbf{A}\mathcal{Y}_{t-1} + \mathbf{C}\varepsilon_t)'] = \mathbf{A}E[\mathcal{Y}_{t-1}\mathcal{Y}_{t-1}']\mathbf{A}' + \mathbf{C}E[\varepsilon_t\varepsilon_t']\mathbf{C}' \\ \mathbf{\Omega} &= \mathbf{A\Omega A}' + \mathbf{C\Sigma C}' \\ \Leftrightarrow \mathrm{vec}(\mathbf{\Omega}) &= [\mathbf{I} - \mathbf{A} \otimes \mathbf{A}]^{-1} (\mathbf{C} \otimes \mathbf{C}) \,\mathrm{vec}(\mathbf{\Sigma}) \end{aligned}$$

Moreover, the auto-covariance matrix is defined as $E[\mathcal{Y}_t\mathcal{Y}_{t-1}']$, thus

$$E[\mathcal{Y}_t\mathcal{Y}_{t-1}'] = E[(\mathbf{A}\mathcal{Y}_{t-1}\mathcal{Y}_{t-1}' + \mathbf{C}\varepsilon_t\mathcal{Y}_{t-1}')] = \mathbf{A}E[\mathcal{Y}_{t-1}\mathcal{Y}_{t-1}'] = \mathbf{A\Omega}.$$

Notice that

$$\mathbf{\Omega} = \begin{bmatrix} E[b_{t+1}^2] & E[b_{t+1}s_t] \\ E[s_tb_{t+1}] & E[s_t^2] \end{bmatrix}, \qquad \mathbf{A\Omega} = \begin{bmatrix} E[b_{t+1}b_t] & E[b_{t+1}s_{t-1}] \\ E[s_tb_t] & E[s_ts_{t-1}] \end{bmatrix}.$$

Recall definitions $\Gamma_b^s \equiv E[b_{t+1}s_t]/E[s_t^2]$ and $\Gamma_b^b \equiv E[b_{t+1}b_t]/E[b_t^2]$ as well as $E[b_{t+1}s_t] = E[s_tb_{t+1}]$, $E[b_{t+1}b_t] = E[b_tb_{t+1}]$, $E[s_{t+1}s_t] = E[s_ts_{t+1}]$, and that $E[s_tu_{t+1}] = E[b_tu_{t+1}] = 0$. Moreover, recall that (6) implies that $E[s_ts_{t+1}] = \phi_bE[s_tb_{t+1}]$ and that $E[b_ts_{t+1}] = \phi_bE[b_tb_{t+1}]$. Thus, we can rewrite (A.1.10), (A.1.11), (A.1.12) and (A.1.13) as

$$\begin{aligned} \psi_v^s(n) =& \mu_{v,1}\Gamma_b^s + \mu_{v,2}\phi_b\Gamma_b^s \\ \psi_p^s(n) =& \mu_{p,1}\Gamma_b^s + \mu_{p,2}\phi_b\Gamma_b^s \\ \psi_v^b(n) =& \mu_{v,1}\Gamma_b^b + \mu_{v,2}\phi_b\Gamma_b^b \\ \psi_p^b(n) =& \mu_{p,1}\Gamma_b^b + \mu_{p,2}\phi_b\Gamma_b^b. \end{aligned}$$

These conditions can be solved for $\psi_v^s(n)$, $\psi_p^s(n)$, $\psi_v^b(n)$, and $\psi_p^b(n)$. In case for $s_b > 0$, this can only be achieved numerically as matrices $\mathbf{A}$ and $\mathbf{C}$ in (A.1.14) also depend on these coefficients.

A.2 Computation of the misspecification equilibrium

Recall the objective (10). Moreover, we have

$$E^s[\mathbf{z}_t^s] = \psi^s(n)s_t, \qquad \text{and} \tag{A.2.1}$$

$$E^b[\mathbf{z}_t^b] = \psi^b(n)b_t. \tag{A.2.2}$$

Thus, we can use (A.1.5) and (A.2.1) to compute

$$(\mathbf{z}_t - E^s[\mathbf{z}_t^s]) = \boldsymbol{\mu}_b b_t + \boldsymbol{\mu}_s s_t + \boldsymbol{\mu}_u u_t - \psi^s(n)s_t.$$

Under the assumption $E[b_t u_t] = E[s_t u_t] = 0$, it follows that

$$E[(\mathbf{z}_t - E^s[\mathbf{z}_t^s])'(\mathbf{z}_t - E^s[\mathbf{z}_t^s])] = E[[b_t'\boldsymbol{\mu}_b' + s_t'\boldsymbol{\mu}_s' + u_t'\boldsymbol{\mu}_u' - s_t'\psi^s(n)']\,[\boldsymbol{\mu}_b b_t + \boldsymbol{\mu}_s s_t + \boldsymbol{\mu}_u u_t - \psi^s(n)s_t]]$$

$$\begin{aligned} E[(\mathbf{z}_t - E^s[\mathbf{z}_t^s])'(\mathbf{z}_t - E^s[\mathbf{z}_t^s])] = {} & (\boldsymbol{\mu}_b'\boldsymbol{\mu}_b)E[b_t^2] \\ & + [\boldsymbol{\mu}_s'\boldsymbol{\mu}_s + \psi^s(n)'\psi^s(n) - \boldsymbol{\mu}_s'\psi^s(n) - \psi^s(n)'\boldsymbol{\mu}_s]E[s_t^2] \\ & + (\boldsymbol{\mu}_u'\boldsymbol{\mu}_u)E[u_t^2] + [\boldsymbol{\mu}_b'\boldsymbol{\mu}_s - \boldsymbol{\mu}_b'\psi^s(n) + \boldsymbol{\mu}_s'\boldsymbol{\mu}_b - \psi^s(n)'\boldsymbol{\mu}_b]E[b_t s_t]. \end{aligned}$$

In consequence, we obtain

$$\begin{aligned} EU^s = -\Big[& (\boldsymbol{\mu}_b'\boldsymbol{\mu}_b)E[b_t^2] + [\boldsymbol{\mu}_s'\boldsymbol{\mu}_s + \psi^s(n)'\psi^s(n) - \boldsymbol{\mu}_s'\psi^s(n) - \psi^s(n)'\boldsymbol{\mu}_s]E[s_t^2] \\ & + (\boldsymbol{\mu}_u'\boldsymbol{\mu}_u)E[u_t^2] + [\boldsymbol{\mu}_b'\boldsymbol{\mu}_s - \boldsymbol{\mu}_b'\psi^s(n) + \boldsymbol{\mu}_s'\boldsymbol{\mu}_b - \psi^s(n)'\boldsymbol{\mu}_b]E[b_t s_t]\Big]. \end{aligned}$$

Likewise, we can use (A.1.5) and (A.2.2) to compute

$$(\mathbf{z}_t - E^b[\mathbf{z}_t^b]) = \boldsymbol{\mu}_b b_t + \boldsymbol{\mu}_s s_t + \boldsymbol{\mu}_u u_t - \psi^b(n)b_t.$$

Therefore it follows that

$$E[(\mathbf{z}_t - E^b[\mathbf{z}_t^b])'(\mathbf{z}_t - E^b[\mathbf{z}_t^b])] = E\big[\big[b_t'\boldsymbol{\mu}_b' + s_t'\boldsymbol{\mu}_s' + u_t'\boldsymbol{\mu}_u' - b_t'\psi^b(n)'\big]\,\big[\boldsymbol{\mu}_b b_t + \boldsymbol{\mu}_s s_t + \boldsymbol{\mu}_u u_t - \psi^b(n)b_t\big]\big].$$

Again we use the assumption $E[b_t u_t] = E[s_t u_t] = 0$ to obtain

$$E[(\mathbf{z}_t - E^b[\mathbf{z}_t^b])'(\mathbf{z}_t - E^b[\mathbf{z}_t^b])] = [\boldsymbol{\mu}_b'\boldsymbol{\mu}_b + \psi^b(n)'\psi^b(n) - \boldsymbol{\mu}_b'\psi^b(n) - \psi^b(n)'\boldsymbol{\mu}_b]E[b_t^2] + (\boldsymbol{\mu}_s'\boldsymbol{\mu}_s)E[s_t^2]$$

$$+ (\boldsymbol{\mu}_u'\boldsymbol{\mu}_u)E[u_t^2] + [\boldsymbol{\mu}_b'\boldsymbol{\mu}_s + \boldsymbol{\mu}_s'\boldsymbol{\mu}_b - \boldsymbol{\mu}_s'\psi^b(n) - \psi^b(n)'\boldsymbol{\mu}_s]E[b_t s_t].$$

In consequence

$$\begin{aligned} EU^b = - \big[& [\boldsymbol{\mu}_b'\boldsymbol{\mu}_b + \psi^b(n)'\psi^b(n) - \boldsymbol{\mu}_b'\psi^b(n) - \psi^b(n)'\boldsymbol{\mu}_b]E[b_t^2] + (\boldsymbol{\mu}_s'\boldsymbol{\mu}_s)E[s_t^2] \\ & + (\boldsymbol{\mu}_u'\boldsymbol{\mu}_u)E[u_t^2] + [\boldsymbol{\mu}_b'\boldsymbol{\mu}_s + \boldsymbol{\mu}_s'\boldsymbol{\mu}_b - \boldsymbol{\mu}_s'\psi^b(n) - \psi^b(n)'\boldsymbol{\mu}_s]E[b_t s_t]\big] . \end{aligned}$$

Finally, one can define $F(n) : [0,1] \to \mathbb{R}$ as $F(n) \equiv EU^s - EU^b$, thus

$$\begin{aligned} F(n) = & \left[\psi^b(n)'\psi^b(n) - \boldsymbol{\mu}_b'\psi^b(n) - \psi^b(n)'\boldsymbol{\mu}_b\right] E[b_t^2] \\ & + [\boldsymbol{\mu}_s'\psi^s(n) + \psi^s(n)'\boldsymbol{\mu}_s - \psi^s(n)'\psi^s(n)] E[s_t^2] \\ & + \left[\psi^s(n)'\boldsymbol{\mu}_b + \boldsymbol{\mu}_b'\psi^s(n) - \psi^b(n)'\boldsymbol{\mu}_s - \boldsymbol{\mu}_s'\psi^b(n)\right] E[b_t s_t]. \end{aligned}$$

B Proofs

B.1 Proof of Proposition B.1.1 (Existence)

Proposition B.1.1 *Let $N_*(\omega) = \{n_* \mid n_* = T_\omega(n_*)\}$ denote the set of misspecification equilibria. As $\omega \to \infty$, N_* has one of the following properties:*

1. *If $F(0) < 0$ and $F(1) < 0$ then $n_* = 0 \in N_*$.*
2. *If $F(0) > 0$ and $F(1) > 0$ then $n_* = 1 \in N_*$.*
3. *If $F(0) < 0$ and $F(1) > 0$ then $\{0, \hat{n}, 1\} \subset N_*$, where $\hat{n} \in (0,1)$ is such that $F(\hat{n}) = 0$.*
4. *If $F(0) > 0$ and $F(1) < 0$ then $n_* = \hat{n} \in N_*$, where $\hat{n} \in (0,1)$ is such that $F(\hat{n}) = 0$.*

Remark B.1.1 *Proposition B.1.1 relies only on the continuity of $F(n)$ and $T_\omega(n)$. If $F(n)$ is monotonic then a stronger statement is possible: Proposition B.1.1 then identifies the full set of misspecification equilibria. In most parameterizations, $F(n)$ is monotonic. When $F(n)$ is non-monotonic it is theoretically possible for there to exist* *multiple interior equilibria, though, in all of the numerical cases examined we found at most 3 misspecification equilibria.*

The proof to Proposition B.1.1 is straightforward, but relies on the existence of a unique restricted perceptions equilibrium for an open set of n. The following Lemma provides the necessary and sufficient conditions for a unique RPE to exist.

Before stating the proposition, note first that the temporary equilibrium equations can be written in the form of an expectational difference equation:

$$X_t = A \begin{bmatrix} b_t \\ s_t \end{bmatrix} + B\hat{E}_t X_{t+1} + C\hat{\epsilon}_t$$

where $X_t' = (s_t, b_t)$ and $\hat{\epsilon}_t$ is a vector of white noise shocks and A, B, C are conformable. Further, denote $EX_tX_t' = \Omega$, $\Gamma_1 = E\begin{bmatrix} b_t \\ s_t \end{bmatrix}\begin{bmatrix} b_{t-1} \\ s_{t-1} \end{bmatrix}'$, and e_j is a (1×2) unit vector with a 1 in the jth element.

Lemma 1 *A unique restricted perceptions equilibrium exists for all n if and only if*

$$\triangle \equiv \det\left(I_4 - P' \otimes B\right) \neq 0$$

where

$$P = \Gamma_1' \left[ne_1' \left(e_1 \Omega e_1'\right)^{-1} e_1 + (1-n)e_2' \left(e_2 \Omega e_2'\right)^{-1} e_2\right]$$

Proof. In an RPE

$$Ee_j \begin{bmatrix} b_{t-1} \\ s_{t-1} \end{bmatrix} \left(X_t - \psi^k e_j \begin{bmatrix} b_{t-1} \\ s_{t-1} \end{bmatrix}\right)' = 0$$

After plugging in for aggregate expectations into the expectational difference equation

$$X_t = \xi \begin{bmatrix} b_t \\ s_t \end{bmatrix} + C\hat{\epsilon}_t$$

where

$$\xi = A + nB\psi^s e_1 + (1-n)B\psi^b e_2$$

Using this notation,

$$
\begin{aligned}
\psi^{k'} &= \left(e_j\Omega e_j'\right)^{-1} Ee_j \begin{bmatrix} b_{t-1} \\ s_{t-1} \end{bmatrix} X_t' \\
&= \left(e_j\Omega e_j'\right)^{-1} e_j\Gamma_1\xi'
\end{aligned}
$$

After plugging in for ψ^s, ψ^b into ξ:

$$
\begin{aligned}
\xi &= A + B\xi\Gamma_1' \left[ne_1' \left(e_1\Omega e_1'\right)^{-1} e_1 + (1-n)e_2' \left(e_2\Omega e_2'\right)^{-1} e_2\right] \\
\Leftrightarrow \xi &= A + B\xi P
\end{aligned}
$$

It follows that

$$
\mathrm{vec}(\xi) = \mathrm{vec}(A) + (P' \otimes B)\,\mathrm{vec}(\xi)
$$

Finally, the RPE coefficient is given by

$$
\mathrm{vec}(\xi) = \left(I_4 - (P' \otimes B)\right)^{-1} \mathrm{vec}(A)
$$

and the stated conditions provides necessary and sufficient conditions for a unique ξ. ■

Proof of Proposition B.1.1.

The existence of a set of fixed points $n_* = T_\omega(n_*)$ follow directly from applying Brouwer's theorem, since $T_\omega : [0,1] \to [0,1]$ and $F(n)$ is continuous provided there exists an RPE. Lemma 1 provides the requisite necessary and sufficient conditions. To complete the proof, we simply require establishing the existence of a unique RPE for an open set of n. This is straightforward as for $n = 0$ or $n = 1$ implies that $\triangle = 0$ and ξ is continuous in n.

■

C Bayesian estimation details

Recall the nonlinear state-space model:

$$
\begin{aligned}
X_t &= g\left(X_{t-1}, \Theta\right) + Q(X_{t-1}, \Theta)\nu_t \\
\mathbb{W}_t &= f(X_t, \upsilon_t),
\end{aligned}
$$

where the state vector is

$$X_t' = \left(b_{t+1}, \pi_t, y_t, v_t, s_t, g_t, u_t, w_t, z_t, n_t, MSE_{st}, MSE_{bt}, \mathrm{vec}\left(\psi_t^s\right), \mathrm{vec}\left(\psi_t^b\right)\right),$$

vec $(\cdot)$ is the vectorization operator, the observation variables are

$$\mathbb{W}_t' = \left(y_t, \pi_t, s_t, b_{t+1}, i_t\right),$$

and the parameter vector is

$$\Theta' = \left(\kappa, \alpha, \phi_\pi, \phi_y, \phi_b, \rho_g, \rho_u, \rho_w, \rho_z, \sigma_g, \sigma_u, \sigma_w, \sigma_z, \omega, \gamma_1, \gamma_2\right).$$

The measurement and state disturbances are υ_t, ν_t respectively. The 4 exogenous shocks follow a linear transition equation with a diagonal matrix whose diagonals are the respective AR(1) coefficients.

In brief, the particle filter, like the Kalman filter, operates in both a prediction and update steps. The first step, given the previous period's particle approximation, is to draw a large number of innovations and then iterate on the state transition equation to yield a predicted next-period state. The predicted particles are then re-weighted based on the most recent data observation, this is the updating step. The updated weights are used to directly compute the likelihood value.

The Bootstrap particle filter, while conceptually straightforward, introduces several computational challenges. First, a stable approximation of the likelihood function requires a large number of particles. Even with efficient parallelization and vectorization each computation of the likelihood approximation is time-consuming. In our application, this is especially true since the state-vector consists of 34 variables. Adopting stochastic gradient learning versus constant gain learning lowers the computational cost considerably. With constant gain least-squares, the particle filter algorithm would require inverting the regressor covariance matrix 10^{12} times. Second, measurement noise in the observation equation is necessary for an accurate particle filter approximation of the likelihood function. The consequence is that the parameter estimates are estimated with greater uncertainty. We follow Herbst and Schorfheide (2015) in defining measurement error so that $\Sigma_\upsilon = 0.25 \times diag\left[Var\left(\mathbb{W}_T\right)\right]$.

To approximate the likelihood function we use the Bootstrap Particle Filter, as developed in Herbst and Schorfheide (2015):

Algorithm 1 *Bootstrap Particle Filter.*

1. Initialization *Draw the initial particles* $\nu_0^j \overset{iid}{\sim} p(X_0)$ *and let* $W_0^j, j = 1, ..., M$.

2. Recursion *For* $t = 1, ..., T$*:*

(a) Forecasting X_t. *Iterate on the state-transition equation:*

$$\hat{X}_t^j = g(X_{t-1}^j, \Theta) + Q(X_{t-1}^j, \Theta)\nu_t^j, \ \nu_t^j \sim F_\nu(\cdot, \Theta)$$

(b) Forecasting $\mathbb{W}_t$. *Define the weights*

$$\hat{w}_t^j = p(\mathbb{W}_t | \hat{X}_t^j, \Theta)$$

where

$$p(\cdot|\cdot) \approx \frac{1}{M} \sum_{j=1}^{M} \hat{w}_t^j W_{t-1}^j$$

and

$$\hat{w}_t^j = (2\pi)^{-n/2} |\Sigma_\upsilon|^{-1/2} \times \exp\left\{ -\frac{1}{2} \left[\mathbb{W}_t - f(\hat{X}_t^j) \right]' \Sigma_\upsilon^{-1} \left[\mathbb{W}_t - f(\hat{X}_t^j) \right] \right\}$$

where Σ_υ *is the covariance matrix for the measurement errors.*

(c) Updating. *Normalize weights:*

$$\hat{W}_t^j = \frac{\hat{w}_t^j W_{t-1}^j}{\frac{1}{M} \sum_{j=1}^{M} \hat{w}_t^j W_{t-1}^j}$$

(d) Selection. *Define* $\hat{m} = M / \left(M^{-1} \sum_{j=1}^{M} (\hat{W}_t^j)^2 \right)$ *and let* r_t *be an indicator variable whenever* $\hat{m}_t < M/2$.

Case 1: $r_t = 1$ *Let* $X_t^j, \ j = 1, ..., M$ *denote* M *iid draws from a multinomial distribution with support points/weights* $\left\{ \hat{X}_t^j, \hat{W}_t^j \right\}$ *and set* $W_t^j = 1, \forall M$.

Case 2: $r_t = 0$ *Then set* $X_t^j = \hat{X}_t^j$ *and* $W_t^j = \hat{W}_t^j$.

3. Likelihood approximation*:*

$$\ln \hat{p}\left(\mathbb{W}^T | \Theta\right) = \sum_{t=1}^{T} \ln \left(M^{-1} \sum_{j=1}^{M} \hat{w}_t^j W_{t-1}^j \right)$$

Finally, we use a Metropolis-Hastings algorithm to construct the posterior distribution:

$$p(\Theta|\mathbb{W}_T) \propto p(\mathbb{W}_T|\Theta) \times p(\Theta)$$

where $p(\Theta)$ is the prior distribution and $p(\Theta|\mathbb{W}_T)$ is the object of interest. Our algorithm samples from the posterior distribution through an adapted Random-Walk Metropolis Hastings (RWMH) MCMC technique, with the particle-filter based estimate of the likelihood function. Convergence properties of the algorithm are discussed in Herbst and Schorfheide (2015). One challenge for the RWMH is identifying a good candidate distribution to draw from, and to be sure that the chain samples from the entire support of the distribution. We do this in two ways. First, we use an adaptive approach to recursively update the candidate distribution combined with a long transient burn-in period. Second, in 99% of the time we draw from a mixture distribution, each proportional to the recursively updated candidate that is centered on the previously accepted draw from the RWMH. The remaining time the draws come naively from the candidate distribution. This ensures that the algorithm is not concentrated near a local maxima and that convergence occurs relatively quickly. We also used a tempering procedure in the early stages of the burn-in period.

We can now describe the Metropolis-Hastings algorithm.

Algorithm 2 *Metropolis-Hastings. For $i = 1, ..., N$*

1. *Draw a candidate θ from $q(\theta|\Theta^{i-1})$*
2. *Set $\Theta^i = \theta$ with probability*

$$\alpha\left(\theta|\Theta^{i-1}\right) = \min\left\{1, \frac{\hat{p}(\mathbb{W}^T|\theta)p(\theta)/q(\theta|\Theta^{i-1})}{\hat{p}(\mathbb{W}^T|\Theta^{i-1})p(\Theta^{i-1})/q(\Theta^{i-1}|\theta)}\right\}$$

where $\hat{p}(\cdot|\cdot)$ is computed using Algorithm 1, and $p(\cdot)$ is the prior.

For the candidate density we use a variant on a random-walk Metropolis Hastings. With probability $\delta \approx 1$ we set

$$q(\cdot|\Theta^{i-1}) = N(\Theta^{i-1}, c\hat{\Sigma}_i)$$

where we use a recursive adaptive algorithm to compute $\hat{\Sigma}$:

$$
\begin{aligned}
\bar{\Theta}_i &= \frac{i+1}{100+i+1}\left[\frac{i}{1+i}\bar{\Theta}_{i-1} + \frac{1}{1+i}\Theta^{i-1}\right] + \frac{100}{100+i+1}\Theta^0 \\
\hat{\Sigma}_i &= \frac{i}{100+i}\left[\frac{i-1}{i}\hat{\Sigma}_{i-1} + \left(i\bar{\Theta}_{i-1}\bar{\Theta}'_{i-1} - (i+1)\bar{\Theta}_i\bar{\Theta}'_i\Theta^{i-1}\Theta^{(i-1)'}\right)/i\right] + \frac{100}{100+i}\hat{\Sigma}_0
\end{aligned}
$$

and with complementary probability

$$q(\cdot|\Theta^{i-1}) = N(\bar{\Theta}_i, c\hat{\Sigma}_i)$$

In the estimation, we set $M = 120,000$ and $N = 200,000$. We use a burn-in period of $80,000$ draws. We ran a tuning run to initialize $\hat{\Sigma}$ during which we also used a tempering schedule. The priors are specified below.

Table A1: Prior distribution of parameters

Parameter	Dist.	Para(1)	Para(2)
Structural parameters			
κ	Beta	0.1	0.2
α	Beta	0.6	0.2
Policy parameters			
ϕ_π	Normal	1.50	0.25
ϕ_y	Normal	0.125	0.05
ϕ_b	Inv. Gamma	0.15	0.05
Exogenous shocks			
ρ_g	Beta	0.50	0.20
ρ_u	Beta	0.50	0.20
ρ_w	Beta	0.50	0.20
ρ_z	Beta	0.50	0.20
$100\sigma_g$	Inv. Gamma	0.01	2.00
$100\sigma_u$	Inv. Gamma	0.01	2.00
$100\sigma_w$	Inv. Gamma	0.01	2.00
$100\sigma_z$	Inv. Gamma	0.01	2.00
Learning parameters			
ω	Gamma	5.00	2.00
γ_1	Beta	0.015	0.015
γ_2	Beta	0.03	0.03

Many of our priors are the same as in Eusepi and Preston (2018) and Herbst and Schorfheide (2015). We set the prior for the intensity of choice in line with survey estimates provided in Branch (2004). Our prior for the gain parameters are informed by previous estimates provided in the literature. The observation equation also includes measurement errros, as discussed in the main text. We set these, following Herbst and Schorfheide (2015), as $\Sigma_v = 0.25 \times diag\left[Var\left(\mathbb{W}^T\right)\right]$.

Supplementary Appendix

Additional model details

This Appendix provides additional details on the model and its derivations. The reader is referred to Woodford (2013) for more complete details. All parameters and variables are explained in the paper above.

Households. Imposing Ricardian beliefs (3) onto the consumption rule (1) leads to a consumption function that satisfies Ricardian equivalence:

$$c_t^i = \sum_{T=t}^{\infty} \beta^{T-t} E_t^i \{(1-\beta)(Y_T - g_T) - \beta\sigma(\beta i_T - \pi_{T+1}),$$

where $g_t = G_t + \bar{c}_t$ is a composite consumption shock.

On the other hand, with non-Ricardian beliefs the path of future surpluses has a direct effect on consumption:

$$\begin{aligned} c_t^i =& \sum_{T=t}^{\infty} \beta^{T-t} E_t^i \{(1-\beta)(Y_T - g_T) - \beta\sigma(\beta i_T - \pi_{T+1})\} \\ &+ (1-\beta) b_t^i + \sum_{T=t}^{\infty} \beta^{T-t} E_t^i \{(1-\beta) s_b(\beta i_T - \pi_T) - s_T\}\}. \end{aligned}$$

Evidently, non-Ricardian beliefs lead households to perceive holdings of government debt as real wealth and a change in the expected path for future surpluses can have a real effect on consumption. In our theory, we posit two forecasting models that, in equilibrium, will differ in whether beliefs are Ricardian or not.

Firms. A firm j that can optimally reset price $p_t^*(j)$ will do so to satisfy the first-order condition

$$p_t^*(j) = (1-\alpha\beta) \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left(E_t^j p_T^{\text{opt}} - p_{t-1} \right),$$

where $E_t^j p_T^{\text{opt}}$ is the perceived optimal price in period T. This condition can be written recursively:

$$\begin{aligned} p_t^*(j) =& (1-\alpha\beta)\left(E_t^j p_t^{\text{opt}} - p_{t-1}\right) + (\alpha\beta) E_t^j p_{t+1}^*(j) + (\alpha\beta)\pi_t, \qquad \text{where} \\ E_t^j p_{t+1}^*(j) \equiv& (1-\alpha\beta)\sum_{T=t}^{\infty}(\alpha\beta)^{T-t}\left(E_t^j p_{T+1}^{\text{opt}} - p_t\right). \end{aligned} \tag{6.1}$$

Temporary equilibrium with heterogeneous beliefs. Recall the consumption function recursion:

$$\begin{aligned} c_t^i &= (1-\beta)b_t^i + (1-\beta)(Y_t - \tau_t) - \beta[\sigma - (1-\beta)s_b]i_t - (1-\beta)s_b\pi_t + \beta\bar{c}_t + \beta E_t^i v_{t+1}^i \\ v_t^i &= (1-\beta)(Y_t - \tau_t) - [\sigma - (1-\beta)s_b](\beta i_t - \pi_t) - (1-\beta)\bar{c}_t + \beta E_t^i v_{t+1}^i. \end{aligned}$$

The latter two can be written as

$$c_t^i = (1-\beta)b_t^i + \bar{c}_t - \sigma\pi_t + v_t^i,$$

which, together with $b_t \equiv \int b_t^i di$, $v_t \equiv \int v_t^i di$ and (8) yields aggregate demand

$$Y_t = \bar{c}_t + G_t + (1-\beta)b_t + v_t - \sigma\pi_t. \tag{6.2}$$

To express the aggregate demand equation in explicit dependence of expectations, we apply (7) and (6.2) to the v_t^i recursion and obtain

$$v_t^i = (1-\beta)v_t + (1-\beta)\beta(b_{t+1} - b_t) - \beta\sigma(i_t - \pi_t) + \beta E_t^i v_{t+1}. \tag{6.3}$$

Averaging over expectations in (6.3) and plugging into (6.2) yields (9). Thus, because the continuation variable v_t^i consists of aggregate variables that are beyond the household's control, and the agents understand their optimal consumption plan and perceived budget constraints, the aggregation result in the main text follows immediately. The ease with which the heterogeneous beliefs aggregate follows from the infinite-horizon learning consumption, which depends on household $i's$ subjective *forecasts* of aggregate variables beyond their control. An example of where aggregation of heterogeneous beliefs is made more difficult by higher-order beliefs is provided by Branch and McGough (2009).

Next, as in Woodford (2013, Section 2.3), in equilibrium the optimal price in this model can be expressed as

$$p_t^{\text{opt}} = p_t + \zeta \left(Y_t - Y_t^n\right) + \mu_t, \tag{6.4}$$

where $\zeta > 0$ is a composite term of structural parameters measuring the output elasticity of a firm's optimal price.[20] The exogenous random variable Y_t^n is the natural level of output that captures exogenous demand shocks and μ_t represents exogenous disturbances to the desired markup over marginal cost.

As the firm's price is a decision variable, it is natural to impose that $E_t^j p_t^{\text{opt}} = p_t^{\text{opt}}$. It follows, then, from plugging (6.4) and (4) into (6.1) that

$$p_t^*(j) = (1-\alpha)p_t^* + (1-\alpha\beta)\left[\zeta y_t + \mu_t\right] + \alpha\beta E_t^j p_{t+1}^*(j).$$

Again averaging across firms, defining the output gap as $y_t \equiv Y_t - Y_t^n$, parameter $\kappa \equiv [(1-\alpha)(1-\alpha\beta)\zeta]/\alpha$, and the cost-push supply shock as $u_t \equiv \{[(1-\alpha)(1-\alpha\beta)]/\alpha\}\mu_t$, yields the New Keynesian Phillips curve

$$\pi_t = (1-\alpha)\beta \int E_t^j p_{t+1}^*(j)dj + \kappa y_t + u_t. \tag{6.5}$$

As for the households, after applying the law of iterated expectations a firm j sets

$$p_t^*(j) = (1-\alpha)p_t^* + (1-\alpha\beta)\left[\zeta y_t + \mu_t\right] + \alpha\beta E_t^j p_{t+1}^*$$

and, an aggregate New Keynesian Phillips Curve results after averaging across all firms:

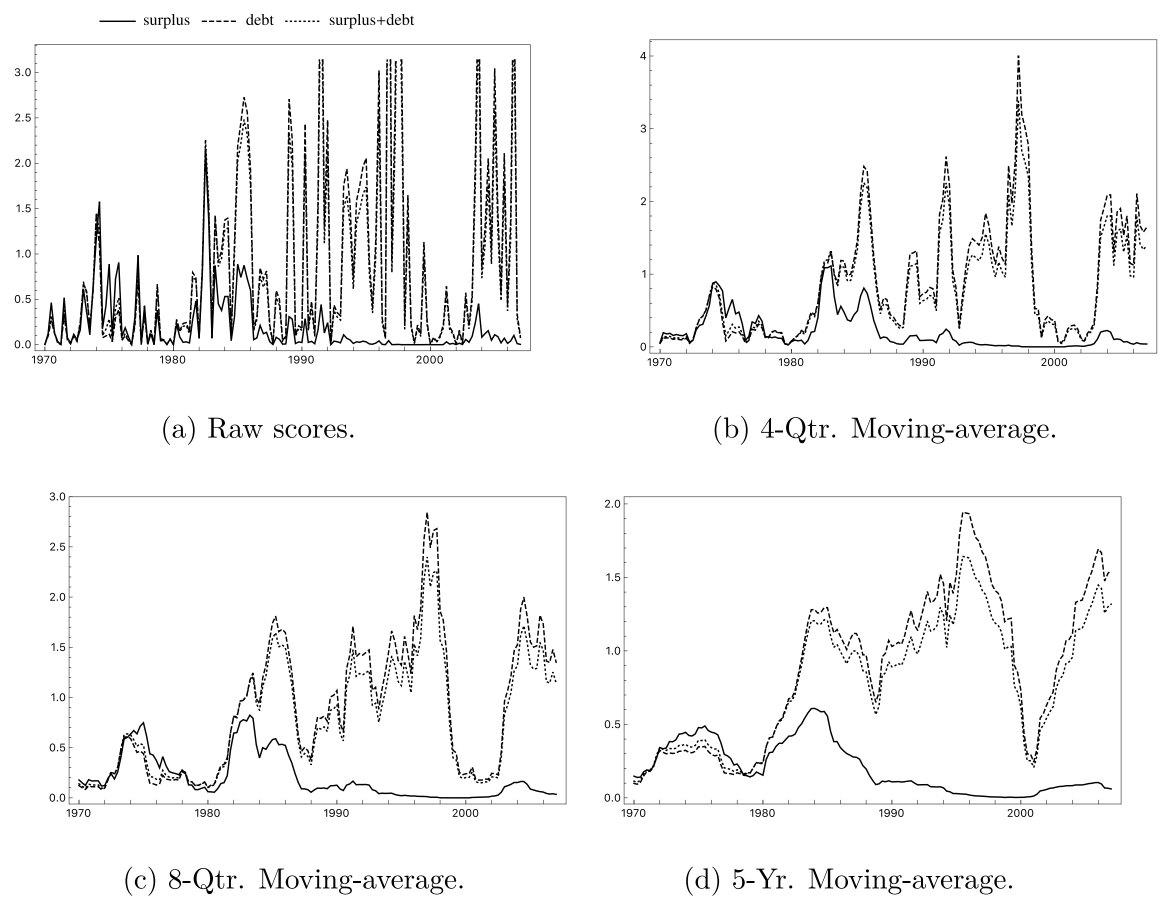
$$\pi_t = (1-\alpha)\,\beta \hat{E}_t p_{t+1}^* + \kappa y_t + u_t.$$

Restricted perceptions vs. rational expectations

Since $n = 0$ is equivalent to the rational expectations model, our estimation results present an econometric test of our model of restricted perceptions against rational expectations. The results tell us that the data prefer the specification of the model with a non-Ricardian equilibrium over most of the sample. To better assess the plausibility of these results, we present (informal) evidence from the Survey of Professional Forecasters (SPF). Although beyond the scope of this paper, a complete empirical analysis

[20] The term is defined in Woodford (2003, ch.3-4).

would use the empirical framework in Branch (2004) to analyze the probability that an individual-level survey forecast was made by a simple model with a restricted set of fiscal variables. As a first pass of providing some external evidence, we compute the statistical scores of the median SPF forecast across three different possible sets of forecasting model regressors: one that includes the primary surplus only, one that includes the debt, and one that includes both. Specifically, we compute moving averages of the statistical score $Ex_{j,t-1}\left(\pi_t - \pi^e_{t-1,t}\right)$ where $x_{j,t} \in \{s_t, b_t\}$, π_t is the PCE inflation rate, and $\pi^e_{t-1,t}$ is the one-step ahead median SPF survey forecast. The results are plotted in Figure A1. Notice that within a restricted perceptions equilibrium, the (time-)average score should be zero. Thus, if the surplus model leads to a lower, and near zero, score then this provides indirect evidence in favor of a restricted perceptions equilibrium with non-Ricardian beliefs. The results in the scores-Figure shows that, beginning in the late 1980's, the median SPF is consistent with a greater share of forecasters using the primary surplus as the fiscal variable. In fact, in the late 1990's that score vector is near zero, as predicted by a non-Ricardian restricted perceptions equilibrium.



(a) Raw scores.

(b) 4-Qtr. Moving-average.

(c) 8-Qtr. Moving-average.

(d) 5-Yr. Moving-average.

Figure A1: Measured scores for different sets of predictors in the Survey of Professional Forecasters. Each panel computes the scores with different moving average lengths. A score close to zero is consistent with a restricted perceptions equilibrium.