Equilibrium Selection, Observability and Backward-stable Solutions*

George W. Evans
University of Oregon and University of St Andrews

Bruce McGough
University of Oregon

February 28, 2018

Abstract

We examine robustness of stability under learning to observability of exogenous shocks. Regardless of observability assumptions, the minimal state variable solution is robustly stable under learning provided the expectational feedback is not both positive and large, while the nonfundamental solution is never robustly stable. Overlapping generations and New Keynesian models are considered and concerns raised in Cochrane (2011, 2017) are addressed.

JEL Classifications: E31; E32; E52; D84; D83

Key Words: Expectations; Learning; Observability; New Keynesian.

1 Introduction

We explore the connection between shock observability and equilibrium selection under learning. Under rational expectations (RE) shock observability does not affect equilibrium outcomes if current aggregates are observable, but stability under adaptive learning is altered. The common assumption that exogenous shocks are observable to forward-looking agents has been questioned by prominent authors, including

---

Cochrane (2009) and Levine et al. (2012). We examine the implications for equilibrium selection of taking the shocks as unobservable.

It is useful first to recall some standard terminology. An equilibrium, viewed here as a stochastic process, is called "explosive" if it is unbounded (almost surely) as time goes to infinity; otherwise it is called "non-explosive." In Section 2 we also refer to non-explosive solutions as "forward-stable" and we define "backward-stable" in an analogous fashion. A linear model is said to be "determinate" if it has a unique non-explosive equilibrium; it is called "indeterminate" if it has multiple non-explosive equilibria; and it is called "explosive" otherwise.

With this terminology our results are easily summarized. Using a simple model, we establish that determinacy implies the existence of a unique rational expectations equilibrium that is robustly stable under learning. Our results also cover the indeterminate case. We then examine the concerns raised in Cochrane (2009, 2011, 2017) about the New Keynesian model. These concerns arise in large part by his adoption of RE as a modeling primitive. We view RE as more naturally arising as an emergent outcome of an adaptive learning process, and we find that by modeling agents as adaptive learners Cochrane's concerns vanish.

To elaborate more fully, the presence of multiple equilibria in rational expectations models has resulted in frequent debates on selection criteria. In the determinate case it is common practice to pick the unique non-explosive RE solution. This choice has theoretical support when the explosive paths can be shown not to be legitimate equilibrium paths based on transversality conditions or on no-Ponzi-game or other constraints. However, there are cases in which explosive RE paths meet all the relevant equilibrium conditions. Furthermore, there are many models that are indeterminate, i.e. have multiple non-explosive RE solutions.

Adaptive learning provides a natural equilibrium selection device. Under adaptive learning agents are assumed to form expectations using forecasting models, which they update over time in response to observed data. There is a well-developed theory that allows the researcher to assess whether agents, using least-squares updating of the coefficients of their forecasting model, will come to behave in a manner that is asymptotically consistent with RE, i.e. whether the rational expectations equilibrium (REE) is stable under learning: see Marcet and Sargent (1989) and Evans and Honkapohja (2001).[1] For the cases of equilibrium multiplicity studied in this paper we adopt stability under adaptive learning as our selection criterion.

Both RE and the adaptive learning approach require precise specification of the information structure. A common assumption is that agents are able to condition time t forecasts on t-dated exogenous shocks. While in some environments this assumption may be quite natural, in others, e.g. for aggregate productivity or taste

---

[1] Using the standard linear three-equation reduced-form set-up, stability under adaptive learning in New Keynesian models was first studied by Bullard and Mitra (2002).

shocks, it may be difficult to defend. These issues have been raised recently in policy-related literature: Cochrane (2009, 2011) has argued that the New Keynesian model's exogenous shocks, specifically the monetary policy shocks, are most naturally taken as unobservable. RE models that exclude from information sets certain contemporaneous variables have been prominent: Lucas (1973) did precisely this in his famous islands model, and Mankiw and Reis (2002) have exploited the same idea in their sticky information DSGE environment. A central part of this paper is to explore the connection between observability and equilibrium selection.

In Section 2 we develop a general theory within the context of a simple forward-looking model with an exogenous shock. We identify two REE of interest: the fundamental, or "minimal state variable" (MSV) solution, which is always non-explosive; and a "non-fundamental" (NF) solution, which may or may not be non-explosive. We then turn to stability under learning. In a model with observable shocks, McCallum (2009a) showed that determinacy implies that only the MSV solution is stable under learning. However, Cochrane (2009, 2011) argued that McCallum's stability results hinged on observability of these shocks. We study stability under learning of the two solutions, taking into account the possibility of unobserved exogenous shocks. The central result of our paper is that, regardless of the observability assumption, the MSV solution is robustly stable under learning, provided only that the positive feedback from expectations is not too large. In contrast the NF solution is never robustly stable under learning.

While our analysis is done in a general framework, the motivation for this study is the collective issues raised by Cochrane (2009, 2011, 2017) in connection with the NK model. To address these issues, Section 3 studies three applications. The first is the flexible-price NK model employed by McCallum (2009a) and Cochrane (2009) in their interchange.[2] We find that provided the interest-rate rule satisfies the Taylor principle the MSV solution is both the unique non-explosive solution and the unique robustly stable REE under learning, regardless of shock observability. In particular, in the NK model adaptive learning selects the REE solution typically adopted by practitioners. The second application is a discrete-time version of the sticky-price NK model used by Cochrane (2017) to study the backward-stable solution under an interest-rate peg. We demonstrate that this solution is precisely the MSV solution and that it is not stable under learning. The last application is an overlapping generations model with production and money. This model can be either determinate (in which case the MSV solution is the unique non-explosive solution) or indeterminate (in which case the NF solution is also non-explosive). We show that in either case the MSV solution is robustly stable under learning and the NF solution is not.

---

[2]McCallum (2012) provides detailed comments on this interchange and discusses further the issue of equilibrium selection.

2 Model and Results

Our key results can be presented most effectively in a linear univariate framework. The model is given by

$$y_t = \beta E_t y_{t+1} + v_t \text{ and } v_t = \rho v_{t-1} + \varepsilon_t, \text{ where } 0 < \rho < 1. \tag{1}$$

Here v_t captures a positively autocorrelated stationary exogenous process, with ε_t zero mean, *iid* and having bounded support. The parameter β measures the expectational feedback in the model. We exclude the non-generic cases in which $|\beta| = 1$ or $\beta\rho = 1$. The expectational feedback parameter β plays an important role in the assessment of equilibrium multiplicity. For convenience we do not include an intercept in (1); thus y should be interpreted as in deviation from mean form.[3] We assume throughout that y_t is in the information set at time t. Under adaptive learning it can matter whether or not v_t is assumed to be observable by agents at time t; under RE, even if v_t is assumed unobservable, agents can deduce at time t its value.

An REE is a stochastic process y_t that satisfies (1). As noted in the introduction a standard selection criterion is that an equilibrium be non-explosive, i.e. "forward stable" in the sense that it remains bounded as $t \to \infty$. Also, we noted there that when the model is determinate there is a unique forward-stable solution, and when the model is indeterminate there are many such solutions. We examine both determinate and indeterminate specifications of our models.

In our investigations, we consider two types of equilibria that are of central focus in the literature: the minimal state variable solution and a non-fundamental or "bubble" solution. These solutions are given by

$$\text{MSV:} \qquad y_t = (1 - \beta\rho)^{-1} v_t \tag{2}$$

$$\text{NF:} \qquad y_t = \beta^{-1} y_{t-1} - \frac{1}{\beta\rho} v_t. \tag{3}$$

Economists frequently adopt the MSV solution as the equilibrium of interest, in part because it is generically forward stable, and in particular is the unique forward-stable solution when the model is determinate, i.e. when $|\beta| < 1$. When the model is indeterminate, i.e. $|\beta| > 1$, the NF solution is also forward stable.[4]

We use adaptive learning to select between these equilibria. The representative agent is assumed to use a forecasting model, usually referred to as a perceived law of motion, or PLM, to form expectations. The PLM depends on parameters that are estimated and updated over time as new data become available. An REE is stable

---

[3] The omission of an intercept is without loss of generality because we include a constant in the agent's forecasting model to ensure robustness in learning the steady state.

[4] While there are many non-fundamental solutions, (3) is uniquely identified as having a finite AR-representation.

under learning if the parameter estimates converge to values that deliver RE forecasts. As we will see the analysis of stability depends on whether v_t is observable.

2.1 Observable shocks

In this section we assume that v_t is observable to learning agents. To assess stability of the MSV and NF solutions, we employ the E-stability principle. In each case we provide agents with a PLM consistent the solution under consideration:

$$\text{MSV PLM: } y_t = a + bv_t, \text{ or NF PLM: } y_t = a + \phi y_{t-1} + bv_t.$$

Let θ denote the PLM parameters, i.e. $\theta = (a, b)$ or $\theta = (a, \phi, b)$. The data generating process implied by the PLM specifies a map $T(\theta)$ taking PLM parameters to their actual law of motion (ALM) counterparts. A fixed point θ^* of this T-map corresponds to an REE. If the differential equation $\dot{\theta} = T(\theta) - \theta$ is Lyapunov stable at θ^* the REE is said to be E-stable. The E-stability principle states that E-stable REE are stable under least-squares and closely related updating rules.

The REE parameters for the MSV and NF solutions are $\theta^* = (0, (1 - \beta\rho)^{-1})$ and $\theta^* = (0, \beta^{-1}, -(\beta\rho)^{-1})$, respectively. Because we have assumed y is in deviation from mean form, learning $a = 0$ corresponds to agents correctly learning the steady state. The MSV PLM implies that $\hat{E}_t y_{t+1} = a + b\rho v_t$, where $\hat{E}_t y_{t+1}$ denotes the conditional expectation of learning agents based on a and b. The corresponding T-map is $T(a, b) = (\beta a, 1 + \beta\rho b)$, and it is easily verified that $\theta^* = (0, (1 - \beta\rho)^{-1})$ is E-stable and hence stable under least-squares learning if $\beta < 1$. The NF solution PLM parameters $(a, \phi, b) = (0, \beta^{-1}, -(\beta\rho)^{-1})$ correspond to an REE. However, observability of y_t and the fact that $\hat{E}_t y_{t+1} = a + \phi y_t + b\rho v_t$ imply the T-map $T(a, \phi, b) = (1 - \beta\phi)^{-1}(\beta a, 0, 1 + \beta\rho b)$ for $\phi \neq \beta^{-1}$. It follows that the NF solution corresponds to a singularity in the T-map at $\phi = \beta^{-1}$. As is well known,[5] this singularity implies the NF solution cannot be E-stable: for (a, b) near $(0, -(\beta\rho)^{-1})$ and $\phi < \beta^{-1}$ close to β^{-1}, the belief parameter ϕ converges to zero and if $\beta < 1$ then $\theta = (a, \phi, b)$ converges to the MSV solution under $\dot{\theta} = T(\theta) - \theta$. The following Theorem, known to the literature, summarizes these results:

Theorem 1 *If the exogenous shocks are observable in the model (1) then (i) the MSV solution is stable under learning when $\beta < 1$ and not stable under learning if $\beta > 1$, and (ii) the NF solution is not stable under learning.*

In fact, the MSV solution is robustly E-stable in the sense that it is locally stable even when the PLM is overparameterized by including any finite number of lags of

---

[5] See Evans (1989), Evans and McGough (2005a) and McCallum (2009a).

5

y_t.[6]

For an application in Section 3 it will be useful to generalize this theorem to the multivariate case. Specifically we consider model

$$y_t = \beta E_t y_{t+1} + \gamma u_t \text{ and } u_t = \Gamma u_{t-1} + \eta_t, \tag{4}$$

with $y \in \mathbb{R}^n$ and $u \in \mathbb{R}^k$. Here η_t is zero mean, *iid* and having bounded support. We assume that the roots of Γ are in $(0, 1)$ and that β is invertible. The corresponding MSV solution is

$$y_t = \psi u_t, \text{ where vec}(\psi) = (I - \Gamma' \otimes \beta)^{-1} \text{vec}(\gamma),$$

and the NF solution is

$$y_t = \beta^{-1} y_{t-1} - \beta^{-1} \gamma \Gamma^{-1} u_t.$$

We have the following theorem:

Theorem 1′ *If the exogenous shocks are observable in the model (4) then (i) the MSV solution is stable under learning when the real parts of the eigenvalues of β are less than one, and not stable under learning if any eigenvalue of β has real part greater than one. (ii) The NF solution is not stable under learning.*

The proof is in the Appendix.

Some observations are warranted. First, since the model is purely forward-looking, invertibility of β is not restrictive in the sense that by a change of variables the model can be reduced to a lower-dimensional system with an associated invertible matrix. Second, in this multivariate case there are nonfundamental solutions in which the coefficient on y_{t-1} is not β^{-1}. These solutions are also unstable under learning but establishing this result requires additional analysis. Finally, while the instability of the analog to NF solutions continues to hold in univariate models with predetermined variables, in a multivariate setting these solutions may be stable under learning for certain parameterizations: see Evans and McGough (2005b) and McGough, B., Meng, Q. and Xue, J. (2013).

2.2 Unobservable shocks

We now assume that v_t is not observable to learning agents. To determine the appropriate PLMs we rewrite the MSV and NF solutions in terms of lags of y_t:

$$\text{MSV:} \quad y_t = \rho y_{t-1} + (1 - \beta \rho)^{-1} \varepsilon_t \tag{5}$$

$$\text{NF:} \quad y_t = (\beta^{-1} + \rho) y_{t-1} - \beta^{-1} \rho y_{t-2} - \frac{1}{\beta \rho} \varepsilon_t. \tag{6}$$

---

[6]If the model is indeterminate there are many forward-stable "sunspot" solutions. If $\beta > 1$ these solutions are not stable under learning; if $\beta < -1$ then there exist PLMs that impart stability under learning to these solutions. See, for example, Evans and McGough (2011).

We consider PLMs allowing for a finite number of lags of y:

$$y_t = a + \sum_{n=1}^{N} \phi_n y_{t-n} + \xi_t, \tag{7}$$

where ξ_t represents the agents' perceived error term, assumed to be a zero mean *iid* stochastic process independent of lagged y. Notice that agents' beliefs are summarized by the vector $\psi_N = (a, \phi)$, where $\phi = (\phi_1, \dots, \phi_N)$.

Given their beliefs, agents form expectations as $\hat{E}_t y_{t+1} = a + \sum_{n=1}^{N} \phi_n y_{t+1-n}$. Again, agents are assumed to use y_t to forecast y_{t+1}. Joining these forecasts to the model (1) leads to the ALM as before. Using $v_t = \rho v_{t-1} + \varepsilon_t$ and letting $\gamma(\phi_1) = (1 - \beta\phi_1)^{-1}$, we have the ALM

$$y_t = \gamma(\phi_1)\beta a(1 - \rho) + \rho y_{t-1} + \gamma(\phi_1)\varepsilon_t, \text{ for } N = 1$$

$$y_t = \gamma(\phi_1)\beta a(1 - \rho) + (\rho + \gamma(\phi_1)\beta\phi_2)y_{t-1} - (\gamma(\phi_1)\beta\rho\phi_2)\, y_{t-2} + \gamma(\phi_1)\varepsilon_t, \text{ for } N = 2$$

$$y_t = \gamma(\phi_1)\beta a(1 - \rho) + (\rho + \gamma(\phi_1)\beta\phi_2)y_{t-1} + \gamma(\phi_1)\beta \sum_{n=2}^{N-1}(\phi_{n+1} - \rho\phi_n)y_{t-n}$$

$$- (\gamma(\phi_1)\beta\rho\phi_N)\, y_{t-N} + \gamma(\phi_1)\varepsilon_t, \text{ for } N \geq 3.$$

The functional forms of the forecasting model and the data generating process exhibit the same linear dependency. This allows us to identify the T-map as usual, which we label as T_N to track the number of lags of y in the PLM. The corresponding E-stability differential equation is

$$\dot{\psi}_N = T_N(\psi_N) - \psi_N. \tag{8}$$

Set $\psi^{MSV} = (0, \rho)$ and $\psi^{NF} = (0, \beta^{-1} + \rho, -\beta^{-1}\rho)$, corresponding to equations (5) and (6). Define

$$\psi_N^{MSV} = (\psi^{MSV}, \underbrace{0, \dots, 0}_{N-1 \text{ terms}}) \text{ for } N \geq 1, \text{ and } \psi_N^{NF} = (\psi^{NF}, \underbrace{0, \dots, 0}_{N-2 \text{ terms}}) \text{ for } N \geq 2.$$

Note that $\psi_1^{MSV} = \psi^{MSV}$ and $\psi_2^{NF} = \psi^{NF}$. The following result characterizes the fixed points of T_N.

Lemma 1 *The unique fixed point of T_1 is ψ^{MSV}. For $N \geq 2$ the fixed points of T_N are given by ψ_N^{MSV} and ψ_N^{NF}.*

The proof of this and all results are in the Appendix. This Lemma shows that there are precisely two equilibrium processes y_t consistent with forecasting models of the form (7), and that they correspond to the MSV and NF solutions.

It is not a priori obvious how many lags agents should include in their forecasting model. To account for this we adopt the following definition of robust stability:

7

Definition 1 *The MSV solution is robustly stable under learning provided ψ_N^{MSV} is a Lyapunov-stable rest point of (8) for all $N \geq 1$. The NF solution is robustly stable under learning provided ψ_N^{NF} is a Lyapunov-stable rest point of (8) for all $N \geq 2$.*

A robustly stable REE continues to be stable under learning when the PLM includes more lags of y than the minimum needed. We have

Theorem 2 *If the exogenous shocks are not observable in the model (1) then (i) the MSV solution is robustly stable under learning when $\beta < 1$ and not robustly stable under learning if $\beta > 1$, and (ii) the NF solution is not robustly stable under learning.*

This shows how the stability results in the observable case extend to the unobservable case.[7] Further, we note that the instability result for the NF solution is even stronger than stated in the theorem: the NF solution fails to be E-stable for each $N \geq 3$.

We remark that when $\beta < 1$ the MSV solution is robustly stable under learning regardless of whether the model is determinate or indeterminate. Conversely, when $\beta > 1$ the MSV (backward-stable) solution is not robustly stable under learning – we return to this case in an example below. The NF solution is explosive when $|\beta| < 1$ and non-explosive when $|\beta| > 1$; however, in both cases the NF solution fails to be robustly stable under learning.

2.3   Discussion: the backward-stable solution

As suggested above, when the model is determinate, forward stability is often adopted as a equilibrium selection criterion. This adoption reflects that forward-stable equilibria typically satisfy certain agent-level optimality conditions and other feasibility constraints. On the other hand, as pointed out by a number of authors, including Cochrane (2009), equilibria that are not bounded as $t \to \infty$ need not necessarily violate these conditions and constraints, particularly if the unbounded variables are nominal. Nonetheless requiring forward stability remains a widely adopted selection criterion, and indeed Cochrane (2017) extends this notion to apply in the indeterminate case by including the requirement that a solution also be "backwards stable," i.e. that the equilibrium remains bounded as $t \to -\infty$.

Here we construct the backward-stable solution in a simple indeterminate model, and we start with the nonstochastic version: $y_t = \beta E_t y_{t+1}$ with $|\beta| > 1$. The collection of perfect foresight solutions takes the form $y_t = C\beta^{-t}$, for any $C \in \mathbb{R}$. Observe that

---

[7]The proof of Theorem 2 is somewhat technical, requiring center-manifold reduction and bifurcations analysis, which makes extensions non-trivial; however, we anticipate that this result holds in the multivariate case. When predetermined variables are introduced, a new equilibrium concept is required, which further complicates matters.

if $C \neq 0$ then $|y_t| \rightarrow \infty$ as $t \rightarrow -\infty$. In this sense a backward-stable solution requires $C = 0$. This REE corresponds to the MSV solution $y_t = 0$.

The concept can easily accommodate the stochastic version of our model (1). Continuing to set $C = 0$, backward iteration provides that

$$y_t = -\sum_{s \geq 1} \beta^{-s} v_{t-s} + \sum_{s \geq 0} \beta^{-s} \xi_{t-s}, \tag{9}$$

where ξ_t is an arbitrary martingale difference sequence (mds) capturing rational forecast errors. If the process ξ_t is *iid* then y_t, as defined by equation (9) is stationary, and thus in a natural sense is both backward and forward stable. Letting $\xi_t = (1 - \beta\rho)^{-1} \varepsilon_t$, the solution (9) reduces to $y_t = (1 - \beta\rho)^{-1} v_t$, i.e. to the MSV solution (2). Because the MSV solution is the "minimal" backward-stable solution, and in particular does not depend on extraneous exogenous variables, it can be viewed as the natural extension of the backward-stable solution to the stochastic case.

3   Applications

We consider three applications: a flexible-price version of the benchmark New Keynesian model; the standard New Keynesian model; and a simple overlapping generations model.

3.1   A flexible-price endowment economy

We first consider the flexible-price environment in which Cochrane and McCallum pursued their debate on whether the NK model uniquely identifies an equilibrium. The general environment is an endowment economy with competitive markets for money, goods, and one-period risk-free claims, flexible prices and infinitely-lived agents. The representative household aims to maximize discounted expected utility with discount factor $0 < \delta < 1$. Expectations are formed against subjective beliefs, and the utility is given by $u(c_t, m_{t-1}) = (1 - \sigma)^{-1} \left(c_t^{1-\sigma} - 1 \right) + \log \left(m_{t-1} \pi_t^{-1} \right)$, with $\sigma > 0$. Here c is consumption, $\pi_t = p_t/p_{t-1}$ is the inflation factor, and $m_{t-1} = M_{t-1}/p_{t-1}$, where M_{t-1} is nominal money holdings at the end of time $t-1$. The budget constraint is $c_t + m_t + b_t = m_{t-1} \pi_t^{-1} + R_{t-1} \pi_t^{-1} b_{t-1} + \mathcal{E}_t + \mathcal{T}_t p_t^{-1}$, where $\mathcal{E}_t$ is the household's time t real endowment of the perishable consumption good, R_{t-1} is the nominal interest rate factor, b_t is the household's real bond holdings in t, and $\mathcal{T}_t$ is nominal monetary injections (which may be positive or negative). For simplicity we assume $\mathcal{E}_t = 1$. The household's first-order conditions are $c_t^{-\sigma} = \delta R_t \hat{E}_t \left(\pi_{t+1}^{-1} c_{t+1}^{-\sigma} \right)$ and $m_t = \delta \left(R_t - 1 \right)^{-1} R_t c_t^{\sigma}$. Finally, we assume households cannot run Ponzi schemes.

There is no public spending and the government does not issue debt; it only prints money and provides nominal injections (transfers). It follows that $b_t = 0$. The

government's real flow constraint is given by $m_t^s = m_{t-1}^s/\pi_t + \mathcal{T}_t/p_t$, which in nominal terms is simply $M_t^s - M_{t-1}^s = \mathcal{T}_t$. Through transfers, the government can choose the nominal money supply M_t^s so that the nominal interest rate satisfies a Taylor rule $R_t = 1 + \varphi_t f(\pi_t)$, where $f(\pi) = (R^* - 1)(\pi/\pi^*)^{\alpha_\pi R^*/(R^*-1)}$. Here φ_t captures a serially correlated policy shock, and is taken to be a stationary AR(1) process in logs, with unit mean. Finally, the inflation target π^* and the interest rate target R^* are assumed to satisfy $\delta R^* = \pi^*$. The model is closed by imposing market clearing.

There are alternative approaches to decision-making under adaptive learning. We adopt the Euler-equation learning approach for its simplicity and because it aligns with the one-step ahead framework, which was the focus of the debate between Mc-Callum and Cochrane.[8] We assume agents have homogeneous expectations. It follows that all agents consume their endowment every period. Assuming that agents take this into account when forecasting future consumption, the Euler equation implies the temporary equilibrium[9] (TE) consumption and money demands

$$c_t = \left(\delta R_t \hat{E}_t \left(\pi_{t+1}^{-1} \right) \right)^{-\frac{1}{\sigma}} \text{ and } M_t = p_t \left((R_t - 1) \hat{E}_t \left(\pi_{t+1}^{-1} \right) \right)^{-1}. \tag{10}$$

Given φ_t, p_{t-1}, and M_{t-1}^s, temporary equilibrium values $(p_t, \pi_t, R_t, c_t, M_t, M_t^s, \mathcal{T}_t)$ are determined by $c_t = 1$, $M_t = M_t^s, M_t^s = \mathcal{T}_t + M_{t-1}^s$, the Taylor rule, and the identity $\pi_t = p_t/p_{t-1}$, where $\hat{E}_t \left(\pi_{t+1}^{-1} \right)$ is allowed to depend on current endogenous variables.

To see how the temporary equilibrium arises, it is helpful to consider a thought experiment in which we start in equilibrium and consider a reduction in expected inflation, or more precisely an increase in $\hat{E}_t \left(\pi_{t+1}^{-1} \right)$. For simplicity suppose that $\hat{E}_t \left(\pi_{t+1}^{-1} \right)$ is predetermined, i.e. independent of current p_t. The increase in the expectations term $\hat{E}_t \left(\pi_{t+1}^{-1} \right)$ acts to reduce both goods demand and money demand, and to raise the supply of saving. The fall in goods demand puts downward pressure on prices p_t (and π_t), and the rise in the supply of savings corresponds to an increase in bond demand, which, because bonds are in zero net supply, puts downward pressure on interest rates.

In the temporary equilibrium the resulting fall in R_t increases consumption demand, eliminating the excess supply of goods. The extent to which p_t and π_t fall is determined by the Taylor rule, which is implemented by changes in M_t^s via transfer payments.[10] To summarize, an exogenous reduction in expected inflation leads to

---

[8]Other possible implementations of adaptive learning include the long-horizon approach of Preston (2005) and versions of the Adam and Marcet (2011) internal rationality approach. In both of these approaches agents fully solve their dynamic program given beliefs. Our implementation can be viewed as a bounded optimality approach, which is shown in Evans and McGough (2017) to be asymptotically optimal. See also Evans and Honkapohja (2006).

[9]The temporary equilibrium concept was originally introduced by Hicks (1939).

[10]Real money balances increase in the temporary equilibrium. Nominal money supply increases provided $\alpha_\pi > R^* - 1$.

lower inflation and interest rates in the temporary equilibrium. While our discussion presumed that $\hat{E}_t\left(\pi_{t+1}^{-1}\right)$ is predetermined, this is not necessary. If $\hat{E}_t\left(\pi_{t+1}^{-1}\right)$ depends on current π_t this introduces additional simultaneity into the temporary equilibrium system, but the central intuition is unchanged and equilibrium is still well-defined.

To complete this example and bring to bear the analysis of Section 2, we follow standard practice under RE and log-linearize the system around the non-stochastic steady state $\bar{c} = 1$ and $\delta R^* = \pi^*$. Log-linearizing the consumption demand equation (10) and the Taylor rule yield $c_t = -\sigma^{-1}\left(R_t - \hat{E}_t\pi_{t+1}\right)$ and $R_t = \alpha_\pi \pi_t + \left(1 - \frac{\delta}{\pi^*}\right)\varphi_t$ where all variables are now written as proportional deviations from means. Because in equilibrium $c_t = 0$, these equations reduce to

$$\pi_t = \frac{1}{\alpha_\pi}\hat{E}_t\pi_{t+1} + v_t, \tag{11}$$

where v_t is the stationary AR(1) process given by $v_t = -\alpha_\pi^{-1}(1 - \delta\left(\pi^*\right)^{-1})\varphi_t$.

Within a linearized framework equation (11) captures the key temporary equilibrium relationship. As usual, under RE this model is determinate provided the Taylor principle $\alpha_\pi > 1$ is satisfied; however, as Cochrane has emphasized, there remain multiple legitimate RE solutions. We use our results to select among the equilibria and thereby connect to the interchange between Cochrane and McCallum. Specifically, McCallum (2007) showed that determinacy implies E-stability of the MSV solution, and McCallum (2009a) established that the NF solution is not E-stable. Cochrane (2011) argues McCallum's results hinge on the observability of exogenous shocks, and that if they are taken as unobserved then the NF solution can be E-stable even when the model is determinate. Theorem 2 dispels this notion, and extends McCallum's result to unobservable shocks. We conclude that stability under adaptive learning does operate as a selection criterion in this model, and that it singles out the MSV solution, i.e. the usual RE solution adopted by proponents of the NK model.

Cochrane (2009, 2011) raises a number of additional concerns about the NK model in the determinate case that can be addressed using the agent-level development above. First, Cochrane has suggested that there is no mechanism in the NK model that pins down prices. Specifically Cochrane (2011, p. 580) writes: "There is no corresponding mechanism to push inflation to the new-Keynesian value [given by the MSV solution]... [S]upply equals demand and consumer optimization hold for any of the alternative paths ... ". He bases this position in part on his assertion (p. 582) that "[t]he equations of the [NK] model do not specify a causal ordering. They are just equilibrium conditions."

Within the perfect-foresight environment, his argument has merit. However, the absence of causality noted by Cochrane results from adopting RE as a primitive rather than viewing it as an emergent outcome of an agent-level learning process.[11] Under

---

[11]Interpreting REE as an emergent outcome is in line with the evolutionary view discussed in

adaptive learning the causal ordering is as follows. At each point in time agents form expectations of key economic variables based on an estimated forecasting model and on observed data; they then form supply and demand schedules based on these expectations and other agent-level variables like wealth. Market clearing gives the temporary equilibrium prices and quantities. In the next period agents revise the coefficients of their forecasting model, e.g. using least-squares updating, and the process continues. This fully specifies a recursive system that defines an equilibrium path under learning. In particular, under adaptive learning, a forecasting model is precisely specified, and given this forecasting model a unique equilibrium path is pinned down. In this way, the adaptive learning approach can resolve the multiplicity problem inherent in the RE version of the model.

Cochrane also suggests that adopting the MSV solution in the NK paradigm requires that the government be interpreted as threatening to "blow up" the economy if that solution is not selected. Cochrane (2011, Appendix B, p. 3) asks: "Is inflation really determined at a given value because for any other value the Fed threatens to take us to a valid but "unlearnable" equilibrium?" The answer is no: inflation is determined in temporary equilibrium, based on expectations that are revised over time in response to observed data. Threats by the Fed are neither made nor needed.

A final issue raised by Cochrane (2009, p. 1111 and 2011, Appendix B, p. 2) concerns econometric identification. Specifically, he correctly points out that in the MSV solution the Taylor-rule parameter α_π is not econometrically identified. He suggests this undermines the adaptive learning approach. However, as also pointed out in McCallum (2012), this misunderstands adaptive learning by private agents as requiring knowledge of the structural parameters. Under adaptive learning agents need to forecast in order to make decisions, but they do not need to know or estimate structural parameters in order to forecast. Instead they make forecasts the same way that time-series econometricians typically do: by estimating least-squares projections of the variables being forecasted on the relevant observables.

Furthermore, the identification issue raised by Cochrane is in any case largely mitigated in models with learning, as shown in Chevillon, Massmann and Mavroeidis (2010) and Christopeit and Massmann (2017). These authors distinguish between the internal forecasting problem and the external estimation problem. Adaptive learning concerns the former problem, and, as we noted above, identification of structural parameters is irrelevant in our setting for agents making forecasts based on observed data. The latter problem is that of an outside econometrician making inferences about structural parameters given the data. If the data is assumed generated in the MSV REE, then α_π is not identified; however, if the data is assumed generated by learning agents, then, strikingly, α_π *is* identified. These authors demonstrate consistency and provide the (possibly) nonstandard asymptotic distribution.

Sargent (2008).

3.2 A sticky-price NK model

We turn to stability under learning of the backward-stable REE in a discrete-time version of the standard NK model considered in Cochrane (2017). The model is

$$y_t = E_t y_{t+1} - \sigma^{-1}(i_t - E_t \pi_{t+1}) \text{ and } \pi_t = \delta E_t \pi_{t+1} + \gamma y_t + u_t, \text{ or} \qquad (12)$$
$$x_t = F E_t x_{t+1} + G^1 u_t + G^2 i_t, \text{ where } x = (y, \pi)',$$

with appropriate F, G^1, G^2, where it is straightforward to show that the eigenvalues of F satisfy $\lambda_1 > 1 > \lambda_2 > 0$. Here u_t is iid, zero mean. In Appendix B we show that the REE associated to a given deterministic interest-rate policy sequence $\{i_t\}$ may be written as

$$(y_t, \pi_t)' = \psi \cdot \left(\sum_{s \geq 0} \lambda_2^s i_{t-1-s}, E_t \sum_{s \geq 0} \lambda_1^{-s} i_{t+1+s}, i_t, u_t \right)' \equiv \psi \cdot w_t, \qquad (13)$$

which is the discrete-time analog of the backward-stable solution of Cochrane (2017).

Under an interest-rate peg, i.e. if $i_t = 0$ for all t, Theorem 1$'$ implies that the REE (13) is not stable under learning.[12] A non-zero interest-rate path introduces a non-homogeneous component to the solution; and, to handle this, and to provide the greatest chance of stability under learning, we assume that agents observe w_t in time t and know ψ; their only uncertainty involves the constant term. Thus agents use the PLM $x_t = A + \psi w_t$, from which it follows that the ALM is given by $x_t = FA + \psi w_t$. The T-map is thus $T(A) = FA$ and the E-stability differential equation is $\dot{A} = (F - I)A$. Because F has a real eigenvalue $\lambda_1 > 1$, the REE $A = 0$ is not E-stable.
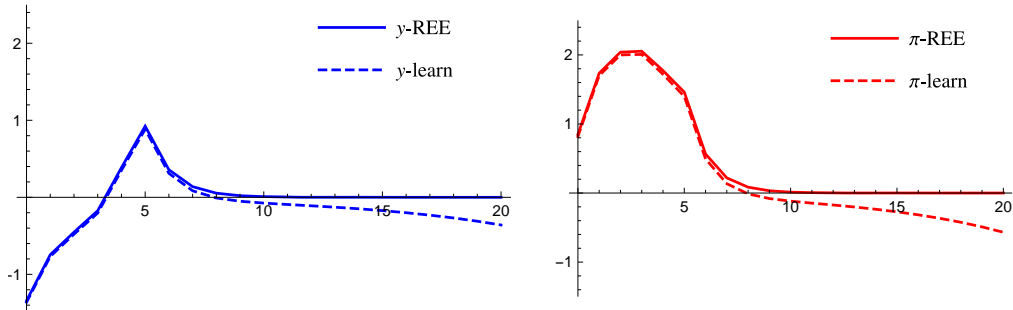

Figure 1: Forward instability of backward-stable solutions

---

[12]This assumes our model (12) fully specifies the economy. The presence of active fiscal policy, and the associated debt dynamics, introduces predetermined variables that may alter determinacy and stability results. This is the subject of current research.

We illustrate the instability using a real-time simulation. To estimate A, agents regress $x_t - \psi w_t$ on a constant. A recursive form of this estimator can be written as

$$A_t = A_{t-1} + \kappa_t \left(x_t - \psi w_t - A_{t-1} \right),$$

where A_t is the estimate of A using data up to time t. Here κ_t is the "gain" sequence, which measures the responsiveness of the estimator to new data. Under "decreasing gain" learning with $\kappa_t = t^{-1}$ this corresponds to the usual least-squares estimator, in which all data are given equal weight, whereas setting $1 > \kappa_t = \kappa > 0$ discounts past data and is referred to as "constant-gain" learning.

We apply this learning algorithm to a central policy experiment studied in Cochrane (2017). Setting the shocks u_t to zero, we assume that, starting at time zero, the interest rises from zero to 2% for 5 periods, before returning to zero. Figure 1 provides both the REE and learning time paths using a gain of $\kappa = 0.1$. This simulation sets initial $A = (-.01, -.01)$ very close to zero, forcefully illustrating the instability of this solution.

We note that Garcia Schmidt and Woodford (2015) take a different line to showing the practical irrelevance of the backward-stable solution. Their analysis uses a "reflective" approach, which is predicated on knowledge by agents of the economic structure and a reasoning process akin to the eductive analysis advocated in Guesnerie (1992). While we regard our analysis and theirs as complementary, we emphasize that our boundedly rational approach does not require that our agents have structural knowledge of the economy.

3.3 Equilibrium selection in the OLG model

In this section, we develop a simple overlapping generations model with productivity shocks as the exogenous stochastic driver. There is a continuum of agents born at each time t indexed by $\omega_t \in \Omega$. Each agent lives two periods, works when young and consumes when old. The population is constant at unit mass. Each agent owns a production technology that is linear in labor and produces a common, perishable consumption good. The technology is subject to idiosyncratic productivity shocks. The agent can sell his produced good in a competitive market for a quantity of fiat currency, anticipating that he will be able to use this currency when old to purchase goods for consumption.

Let $\omega_t \in \Omega$ be the index of a representative agent born in time t. This agent's problem is given by

$$\max \hat{E}(\omega_t) \left(u(c_{t+1}(\omega_t)) - \nu(n_t(\omega_t)) \right) \tag{14}$$

subject to $z_t(\omega_t) n_t(\omega_t) = q_t M_t(\omega_t)$ and $c_{t+1}(\omega_t) = q_{t+1} M_t(\omega_t)$. Here, $n_t(\omega_t)$ is the agent's labor supply when young, $z_t(\omega_t)$ is a positive productivity shock, and $z_t(\omega_t) n_t(\omega_t)$ is his output. Also, q_t is the time t goods price of money and $c_{t+1}(\omega_t)$ is the agent's

planned consumption when old. The expectations operator $\hat{E}(\omega_t)(\cdot)$ denotes the expectation of agent ω_t at time t, taken with respect to his subjective beliefs conditional on the information available to him. This information includes $n_t(\omega_t)$, $M_t(\omega_t)$, $z_t(\omega_t)$ and current and lagged values of q_t.

Assuming CRRA utility with risk aversion σ, and that $\nu' = 1$, we obtain agent ω_t's decision rules:

$$n_t(\omega_t) = \left(\left(\frac{z_t(\omega_t)}{q_t} \right)^{1-\sigma} \hat{E}(\omega_t) \left(q_{t+1}^{1-\sigma} \right) \right)^{\frac{1}{\sigma}} \text{ and } M_t(\omega_t) = \left(\frac{z_t(\omega_t)}{q_t} \hat{E}(\omega_t) \left(q_{t+1}^{1-\sigma} \right) \right)^{\frac{1}{\sigma}}.$$

We note that the quantity of money demanded by agent ω_t at time t, depends on, among other things, the price at time t. Assuming a constant (unit) supply of money, the market-clearing condition $\int_\Omega M_t(\omega_t) d\omega_t = 1$ yields

$$q_t = \left(\int_\Omega \left(z_t(\omega_t) \hat{E}(\omega_t) \left(q_{t+1}^{1-\sigma} \right) \right)^{\frac{1}{\sigma}} d\omega_t \right)^\sigma. \tag{15}$$

which characterizes the equilibrium price path.

To address the observability issue we assume that the productivity shocks $z_t(\omega_t)$ have a common unobserved aggregate component z_t, which in log form is taken to be a stationary AR(1) process and white-noise idiosyncratic components that are independent across agents. Assuming that all agents have the same forecast of aggregates, Appendix C shows that to first order, equation (15) simplifies to

$$\log q_t = (1 - \sigma)\hat{E}_t \log q_{t+1} + \log z_t.$$

This reduces to our model (1) with $y_t = \log q_t$ and $v_t = \log z_t$. This model is indeterminate and has both MSV and NF solutions as stationary equilibria when $\sigma > 2$. Theorem 2 shows that the MSV solution, but not the NF solution, is robustly stable under learning. When $\sigma \in (0, 2)$ the model is determinate. The MSV solution is stationary and robustly stable. The NF solutions diverge away from the steady state and converges either to infinity or zero almost surely. Whether the explosive path represents a genuine equilibrium depends on model details, but in any case the NF solution is not robustly stable.[13]

4   Conclusion

We use adaptive learning to address equilibrium selection, taking into account issues of observability. The key results in our set-up are that provided there is not positive

---

[13]In a related nonstochastic endowment economy Lucas (1986) showed that the path to autarky (corresponding here to $q_t \to 0$) was not stable under a simple learning rule.

expectational feedback that is too large, MSV solutions are robustly stable under learning, while nonfundamental solutions are not robustly stable under learning under any circumstances. In NK models these results indicate the desirability of the Taylor principle, and the dangers of an interest-rate peg, and that when the Taylor principle is followed, the standard solution to the NK model is the appropriate choice.

Appendix A: Proofs of Results

Proof of Theorem 1′. Part (i) follows from Proposition 10.3, p. 238 of Evans and Honkapohja (2001). To establish part (ii), note that the T-map takes the form

$$T(a, \phi, b) = \left((I - \beta\phi)^{-1}, 0, (I - \beta\phi)^{-1} (\gamma + \beta b \Gamma) \right),$$

when $\phi \neq \beta^{-1}$. Under the E-stability dynamics we have in particular that $\dot{\phi} = -\phi$. If $\phi(0) = (1 - \mu)\beta^{-1}$ for $\mu > 0$ small, then the E-stability dynamics are well-defined and imply that $\phi \to 0$. It follows that $\phi = \beta^{-1}$ is not locally E-stable. ∎

Proof of Lemma 1. Recall that with $\gamma(\phi_1) = (1 - \beta\phi_1)^{-1}$ the map T_1 is given by $a \to ((1 - \rho)\gamma(\phi_1)\beta) a$ and $\phi_1 \to \rho$, and the map T_2 is given by $a \to ((1 - \rho)\gamma(\phi_1)\beta) a$, $\phi_1 \to \rho + \beta\gamma(\phi_1)\phi_2$ and $\phi_2 \to -(\gamma(\phi_1)\beta\rho\phi_2)$. For $N \geq 3$, the map T_N is given by

$$
\begin{align}
a & \to ((1 - \rho)\gamma(\phi_1)\beta) a \tag{16} \\
\phi_1 & \to \rho + \beta\gamma(\phi_1)\phi_2 \tag{17} \\
\phi_n & \to \beta\gamma(\phi_1)(\phi_{n+1} - \rho\phi_n), \quad \text{for} \quad n = 2, \ldots, N - 1 \tag{18} \\
\phi_N & \to -(\beta\rho\gamma(\phi_1)) \phi_N. \tag{19}
\end{align}
$$

Direct computation shows that ψ^{MSV} is the unique fixed point of T_1 and that ψ_N^{MSV} is a fixed point of T_N for $N \geq 2$. It is similarly straightforward to verify that ψ_N^{NF} is a fixed point of T_N for $N \geq 2$.

Now let $\psi^* = (a^*, \phi^*)$ be any fixed point of T_N for $N \geq 2$, and assume $\psi^* \neq \psi_N^{MSV}$. Notice that if $\phi_2^* = 0$ then $\phi_1^* \neq 0$. It follows that there is a smallest $\tilde{n} \in \{1, \ldots, N\}$ so that $\phi_{\tilde{n}}^* \neq 0$. Further, $\tilde{n} \geq 2$: indeed $\tilde{n} = 1$ implies $\psi^* = \psi_N^{MSV}$.

We claim that $-\beta\rho\gamma(\phi_1^*) = 1$. Indeed, if $\tilde{n} = N$ the we are done by (19). If $\tilde{n} < N$ then, since $\tilde{n} \geq 2$ and since $\phi_{\tilde{n}+1}^* = 0$, the claim follows by (18). We conclude that $\phi_1^* = \beta^{-1} + \rho$. Then by (17), $\phi_2^* = (\beta\gamma(\phi_1^*))^{-1} (\phi_1^* - \rho) = -\rho/\beta$. Finally, if $N \geq 3$ we may recursively compute, for $n \geq 3$, that $\phi_n^* = ((\beta\gamma(\phi_1^*))^{-1} + \rho) \phi_{n-1}^* = 0$. It follows that $\psi^* = \psi_N^{NF}$, completing the proof. ∎

Proof of Theorem 2. We need to assess E-stability of the MSV and NF solutions for each N using (8) by obtaining the derivative matrix of T_N evaluated at ψ_N^{MSV} and at ψ_N^{NF}. Let T_N^a denote the map (16) and T_N^ϕ denote the map given by (17)-(19). Because the T-map decouples, E-stability can be analyzed by focusing on $\frac{\partial}{\partial a} T_N^a$ and $\frac{\partial}{\partial \phi} T_N^\phi$, separately.

We first consider ψ_N^{MSV}. Evaluated at this fixed point, we have $\partial T_N^a / \partial a = (1 - \beta\rho)^{-1} (1 - \rho)\beta$. Also, $\partial T_1^\phi / \partial \phi_1 = 0$ and, for $N \geq 2$,

$$
\left(\frac{\partial}{\partial \phi'} T_N^\phi \right)_{ij} = \begin{cases} (\beta\rho - 1)^{-1} \beta\rho & \text{if } i = j > 1 \\ (1 - \beta\rho)^{-1} \beta & \text{if } i = j - 1 \\ 0 & \text{else} \end{cases},
$$

which yields eigenvalues of 0 and $(\beta\rho - 1)^{-1}\beta\rho$. For E-stability we need $(\beta\rho - 1)^{-1}\beta\rho < 1$ and $(1 - \beta\rho)^{-1}(1-\rho)\beta\rho < 1$. Since $0 < \rho < 1$ it is straightforward to verify that E-stability holds if and only if $\beta < 1$. This completes the proof of Part 1.

Turning to ψ_N^{NF}, we have $\partial T_N^a/\partial a = -\rho^{-1}(1-\rho)$, which satisfies $\partial T_N^a/\partial a < 1$ since $0 < \rho < 1$. We next compute

$$\left(\frac{\partial}{\partial\phi'}T_N^\phi - I_N\right)_{ij} = \begin{cases} -(\beta\rho)^{-1} - 1 & \text{if } i = j = 1 \\ \beta^{-1} & \text{if } i = j + 1 = 2 \\ -\rho^{-1} & \text{if } i = j - 1 \\ 0 & \text{else} \end{cases}, \tag{20}$$

where I_N is the $N \times N$ identity matrix. The eigenvalues of (20) are -1 and $-(\beta\rho)^{-1}$, plus a zero eigenvalue with multiplicity $N - 2$. E-stability requires that they be less than or equal to zero. It follows that if $\beta < 0$ then ψ_N^{NF} is not E-stable for any $N \geq 2$, and again the result follows. We thus now assume that $\beta > 0$.

Returning to the case $N = 2$, we conclude that with $\beta > 0$ the fixed point ψ_2^{NF} is E-stable. If instead $N > 2$ then the presence of zero eigenvalues implies that the dynamic system (8) is non-hyperbolic at the fixed point ψ_N^{NF}, and higher-order stability analysis is required.

To assess the stability ψ_N^{NF} for $N \geq 3$, we first study the case $N = 3$; the remaining cases will be addressed by induction. For notional simplicity, let ϕ^* be the lag coefficients ψ_3^{NF}. We begin by changing coordinates: $\varphi = \phi - \phi^*$. Define $F : \mathbb{R}^3 \to \mathbb{R}^3$ by $F(\varphi) = T_3^\phi(\varphi + \phi^*) - (\varphi + \phi^*)$. Because F is obtained from $T_3^\phi - I_3$ by an affine transform, the stability of ϕ^* as a fixed point of the system $d\phi/d\tau = T_3^\phi(\phi) - \phi$ is equivalent to the stability of the origin as a fixed point of $d\varphi/d\tau = F(\varphi)$. Thus we study the latter system.

By Taylor's theorem, we may write $F(\varphi) = DF(0)\varphi + \hat{H}(\varphi)$, where $\hat{H}(\varphi) \equiv F(\varphi) - DF(0)\varphi$ is zero to first order. Diagonalize the derivative of F as $DF(0) = S\Lambda S^{-1}$, where $\Lambda = -1 \oplus -(\beta\rho)^{-1} \oplus 0$. Change coordinates: $\zeta = S^{-1}\varphi$. It follows that

$$\begin{aligned} d\zeta/d\tau &= S^{-1}F(\varphi) = S^{-1}DF(0)\varphi + S^{-1}\hat{H}(\varphi) = \Lambda\zeta + S^{-1}\hat{H}(S\zeta), \text{ or} \\ d\zeta/d\tau &\equiv \Lambda\zeta + H(\zeta), \end{aligned} \tag{21}$$

where the last equality defines notation. The stability of the origin as a fixed point of $d\varphi/d\tau = F(\varphi)$ is equivalent to the stability of the origin as a fixed point of (21).

The system (21) has two stable eigenvalues and one zero eigenvalue. By the center-manifold theorem the origin in $\mathbb{R}^3$ is a Lyapunov-stable steady state of (21) if and only if the origin in $\mathbb{R}$ is an Lyapunov-stable steady state of

$$d\zeta_3/d\tau = H^3\left(h_1(\zeta_3), h_2(\zeta_3), \zeta_3\right), \tag{22}$$

where a superscript indicates the coordinate in the range, i.e. $H = (H^1, H^2, H^3)'$, and where functions h_i are zero to first order: $h_i(0) = h_i'(0) = 0$ for $i = 1, 2$.

18

We are required now to determine the explicit form of H^3. Since the last row of S^{-1} is $(0, 0, 1)$, we have that $d\zeta_3/d\tau = F^3(\varphi) = F^3(S\zeta)$. Since the last row of Λ is $(0, 0, 0)$, by equation (21) we have that $d\zeta^3/d\tau = H^3(\zeta)$. It follows that $H^3(\zeta) = F^3(S\zeta)$. We compute the third coordinate of F as

$$F^3(\varphi) = \quad = \quad -(1 - \beta(\varphi_1 + \phi_1^*))^{-1}\beta\rho\varphi_3 - \varphi_3 = -(\varphi_1 + \rho)^{-1}\varphi_1\varphi_3,$$

where the second equality uses $\phi_1^* = \beta^{-1} + \rho$. Continuing, we compute

$$H^3(\zeta) \quad = \quad \left(\beta(\zeta_3 - \rho\zeta_1) - \zeta_2 + \rho^2\right)^{-1}\zeta_3(\zeta_2 - \beta(\zeta_3 - \rho\zeta_1)). \tag{23}$$

We conclude that

$$d\zeta_3/d\tau = \left(\beta(\zeta_3 - \rho h_1(\zeta_3)) - h_2(\zeta_3) + \rho^2\right)^{-1}\zeta_3(h_2(\zeta_3) - \beta(\zeta_3 - \rho h_1(\zeta_3))) \equiv G(\zeta_3),$$

where the second equality defines notation.

Writing $G(\zeta_3) = H^3\left(h_1(\zeta_3), h_2(\zeta_3), \zeta_3\right)$ we have $G' = H_1^3 h_1' + H_2^3 h_2' + H_3^3$ and

$$\begin{aligned}
G'' \quad &= \quad (H_{11}^3 h_1' + H_{12}^3 h_2' + H_{13}^3)h_1' + H_1^3 h_1'' + (H_{21}^3 h_1' + H_{22}^3 h_2' + H_{23}^3)h_1' + H_2^3 h_2'' \\
&+ \quad H_{31}^3 h_1' + H_{32}^3 h_2' + H_{33}^3,
\end{aligned}$$

where the arguments have been suppressed to ease notation, and, for example, $H_{ij}^3 = \partial^2 H^3/\partial\zeta_i\partial\zeta_j$. Using (23) it is immediate the $H_i^3 = 0$ for $i = 1, 2, 3$. Since $h_i = h_i' = 0$ for $i = 1, 2$, it follows, then, that $G' = 0$ and $G'' = H_{33}^3$. Again using (23), we compute directly that $G'' = -2\rho^{-2}\beta$.

The above computations allow us to use Taylor's theorem to write $d\zeta_3/d\tau = -\rho^{-2}\beta\zeta_3^2 + \mathcal{O}(|\zeta_3|^3)$. It follows that the origin is not a Lyapunov stable fixed point of (22): there exists $\varepsilon, \delta > 0$ so that if the system (22) is initialized in the interval $(-\delta, 0)$ the corresponding trajectory will exit the interval $(-\varepsilon, \varepsilon)$ in finite time. We conclude that ψ_3^{NF} is not a Lyapunov-stable fixed point of (8).

Having showing instability for $N = 3$, we now proceed by induction. Some more notation is needed. Denote by Δ the metric on $\mathbb{R}^n$ induced by the max-norm: $\|x\| = \max|x_i|$. Next, recalling the observation above that we may ignore the constant term in the PLM, we denote by $\phi(N) \in \mathbb{R}^N$ a beliefs parameter corresponding to a PLM with no constant term and N lags, and let $\phi^*(N)$ be the lag coefficients of the fixed point ψ_N^{NF}. Let $B_N(\varepsilon)$ be the ε-ball with respect to the metric Δ, centered at $\phi^*(N)$. Finally, let $\phi^t(N)$ be the beliefs time-path implied by $d\phi(N)/d\tau = T_N^\phi(\phi(N)) - \phi(N)$, corresponding to the initial condition $\phi^0(N)$.

Assume instability for the $N - 1$ dimensional system. Let $\varepsilon > 0$ be such that for all $0 < \delta < \varepsilon$, there exists a point $\phi^0(N - 1) \in B_{N-1}(\delta)$ such that $\phi^t(N - 1)$ eventually exits $B_{N-1}(\varepsilon)$. Now let $\phi^0(N) = ((\phi^0(N - 1)), 0) \in B_N(\delta)$. It suffices to show that $\phi^t(N)$ eventually exits $B_N(\varepsilon)$. By looking back at the definition of

19

the T-map, notice that $d\phi_N(N)/d\tau = (-\beta\rho\gamma(\phi_1(N)) - 1)\phi_N(N)$. Since the initial condition has $\phi_N^0(N) = 0$, this differential equation implies that $\phi_N^t(N) = 0$ for all t. Now consider again the T-map, this time focusing on the implied dynamics of ϕ_{N-1}:

$$
\begin{aligned}
d\phi_{N-1}(N)/d\tau &= \beta\gamma(\phi_1(N))(\phi_N(N) - \rho\phi_{N-1}(N)) - \phi_{N-1}(N) \\
&= (-\beta\rho\gamma(\phi_1(N)) - 1)\phi_{N-1}(N);
\end{aligned}
$$

the second equality obtains from the fact that $\phi_N^t(N) = 0$. But this is precisely the differential equation determining the path $\phi_{N-1}^t(N-1)$. From the recursive construction of the T-map, (17)-(18), it further follows that, for $n < N - 1$, the differential equation determining the path of $\phi_n^t(N-1)$ is functionally identical to the differential equation governing the dynamics of $\phi_n^t(N)$. We conclude that if $\phi^0(N) = (\phi^0(N-1), 0)$ then for $n \le N - 1$, $\phi_n^t(N) = \phi_n^t(N-1)$. By the induction hypothesis, there exists a T and an $n \in \{1, \ldots, N-1\}$ so that $|\phi_n^T(N-1)| > \varepsilon$; since $\phi_n^T(N) = \phi_n^T(N-1)$, the proof is complete. ∎

Appendix B: Sticky-price NK Model

Recall that we write the model as $Hx_t = FE_t x_{t+1} + G^1 u_t + G^2 i_t$, where $x = (y, \pi)'$. Factor as follows: $F^{-1}H = S(\lambda_1 \oplus \lambda_2)S^{-1}$, where generically $\lambda_1 > 1 > \lambda_2 > 0$. Changing coordinates to $z = S^{-1}x$ gives

$$
z_{1t} = \theta_1 u_t + \theta_2 i_t + \theta_2 E_t \mathcal{I}_{t+1}^F \text{ and } z_{2t} = C\lambda_2^t + \phi_1 \mathcal{U}_{t-1}^P + \phi_2 \mathcal{I}_{t-1}^P + \mathcal{F}_t^P.
$$

Here ξ_t is an mds capturing the forecast errors associated with z_t, and the remaining terms are defined as follows:

$$
\mathcal{I}_t^F = \sum_{s\ge0} \lambda_1^{-s} i_{t+s}, \mathcal{I}_t^P = \sum_{s\ge0} \lambda_2^s i_{t-s}, \ \mathcal{U}_t^P = \sum_{s\ge0} \lambda_2^s u_{t-s}, \ \mathcal{F}_t^P = \sum_{s\ge0} \lambda_2^s \xi_{t-s}
$$

$$
\phi_1 = \left(-S^{-1}F^{-1}G^1\right)_2, \phi_2 = \left(-S^{-1}F^{-1}G^2\right)_2, \ \theta_1 = \left(S^{-1}H^{-1}G^1\right)_1, \theta_2 = \left(S^{-1}H^{-1}G^2\right)_1.
$$

Letting $\xi_t = -\frac{\phi_1}{\lambda_2} u_t$ yields $\phi_1 \mathcal{U}_{t-1}^P + \mathcal{F}_t^P = -\frac{\phi_1}{\lambda_2} u_t$. Finally, letting

$$
\psi = S \cdot \begin{pmatrix} 1 & \phi_2 & 0 & 0 & -\frac{\phi_1}{\lambda_2} \\ 0 & 0 & \theta_2 & \theta_2 & \theta_1 \end{pmatrix}
$$

the REE associated to a given policy $\{i_t\}$ and shock path $\{u_t\}$ may be written

$$
(y_t, \pi_t)' = \psi \cdot \left(\mathcal{I}_{t-1}^P, E_t\mathcal{I}_{t+1}^F, i_t, u_t\right)' \equiv \psi \cdot w_t,
$$

where here we have set $C = 0$. This equation corresponds to (13).

Appendix C: The OLG Model

Agent ω_t's productivity shock includes both an idiosyncratic and an aggregate component, i.e.

$$
\log(z_t(\omega_t)) = \log(z_t) + \log(\zeta_t(\omega_t)) \text{ and } \log(z_t) = \rho\log(z_{t-1}) + \varepsilon_t,
$$

where the $\log\left(\zeta_t(\omega_t)\right)$ are *iid* mean zero and independent across agents, ε_t is *iid* zero mean with small support and $\log\left(\zeta_t(\omega_t)\right)$ and ε_t are independent processes. The market-clearing condition $\int_\Omega M_t(\omega_t)d\omega_t = 1$ yields

$$q_t = z_t \left(\int_\Omega \left(\zeta_t(\omega_t)\hat{E}(\omega_t)\left(q_{t+1}^{1-\sigma}\right) \right)^{\frac{1}{\sigma}} d\omega_t \right)^\sigma. \qquad (24)$$

Equation (24) characterizes the equilibrium price path. Clearly $z_t(\omega_t)$ can be useful for forecasting q_{t+1} only through its correlation with z_t. Furthermore if the variance of the idiosyncratic shock is large relative the variance of the aggregate shock then the information content of $z_t(\omega_t)$ for forecasting z_{t+1} is small.

An REE is characterized by $q_t = \vartheta z_t E_t q_{t+1}^{1-\sigma}$, where $\vartheta^{\frac{1}{\sigma}} = \int_\Omega \zeta_t(\omega_t)^{\frac{1}{\sigma}} d\omega_t$. However, it is neither natural nor necessary to assume agents observe z_t when making rational forecasts. The required information under RE is encoded in the current price level q_t. Thus under RE agent ω_t has no reason to condition his forecast on $z_t(\omega_t)$. We assume this behavior extends outside of RE as well.

Assuming all agents hold the same forecast of (functions of) aggregates, i.e. $\hat{E}(\omega_t)f(q_{t+1}) = \hat{E}_t f(q_{t+1})$ for all ω_t and well-behaved functions f, the unique, monetary, non-stochastic steady state is given by $q_t = 1$. Provided that the expectations operators of agents are reasonably well-behaved (e.g. are linear, respect first-order approximations, fix constants, etc.), we may linearize (24) around this steady state, and simplify to obtain $\log q_t = (1-\sigma)\hat{E}_t \log q_{t+1} + \log z_t$, since, by independence, $\int_\Omega \log z_t(\omega_t)d\omega_t = \log z_t$.

References

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

[2] Bullard, J. and Mitra, K. (2002), Learning about monetary policy rules. Journal of Monetary Economics 49, 1105-1129.

[3] Chevillon, G., Massmann M. and Mavroeidis S. (2010), Inference in models with adaptive learning. Journal of Monetary Economics 57, 341-351.

[4] Christopeit, N. and Massmann, M. (2017), Estimating structural parameters in regression models with adaptive learning. Econometric Theory, forthcoming.

[5] Cochrane, J. (2009), Can learnability save new-Keynesian models? Journal of Monetary Economics 56, 1109-1113.

[6] Cochrane, J. (2011), Determinacy and identification with Taylor rules. Journal of Political Economy 119, 565-615.

[7] Cochrane, J. (2017), The new-Keynesian liquidity trap. Journal of Monetary Economics, forthcoming.

[8] Evans, G. (1989), The fragility of sunspots and bubbles. Journal of Monetary Economics 23, 297-317.

[9] Evans, G. and Honkapohja, S. (2001), Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.

[10] Evans, G. and Honkapohja, S. (2006), Monetary policy, expectations and commitment. Scandinavian Journal of Economics 108, 15-38.

[11] Evans, G. and McGough, B. (2005a), Indeterminacy and the stability puzzle in nonconvex economies. Contributions to Macroeconomics 5, Issue 1, Article 8.

[12] Evans, G. and McGough, B. (2005b), Stable sunspot solutions in models with predetermined variables. Journal of Economic Dynamics and Control 29, 601-625.

[13] Evans, G. and McGough, B. (2011), Representations and sunspot stability. Macroeconomic Dynamics 15, 80-92.

[14] Evans, G. and McGough, B. (2017), Learning to optimize. Mimeo.

[15] Garcia-Schmidt, M. and Woodford, M. (2015), Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis. NBER Working Paper No. 21614

[16] Hicks, J., (1939), Value and Capital. Oxford University Press, Oxford.

[17] Levine, P., Pearlman, J., Perendia, G. and Yang, B. (2012), Endogenous persistence in an estimated DSGE model under imperfect information. Economic Journal 122, 1287-1312.

[18] Lucas, Jr., R.E. (1973), Some international evidence on output-inflation trade-offs. American Economic Review, 63, 326-334.

[19] Lucas Jr., R.E. (1986), Adaptive behavior and economic theory. Journal of Business 59, S401-S426.

[20] Mankiw, N.G. and Reis, R. (2002), Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. Quarterly Journal of Economics 117, 1295–1328.

[21] Marcet, A. and Sargent, T. (1989), Convergence of least-squares learning mechanisms in self-referential linear stochastic models. Journal of Economic Theory 48, 337-368.

[22] McCallum, B. (2007), E-Stability vis-a-vis determinacy results for a broad class of linear rational expectations models. Journal of Economic Dynamics and Control 31, 1376-1391.

[23] McCallum, B. (2009a), Inflation determination with Taylor rules: is new-Keynesian analysis critically flawed? Journal of Monetary Economics 56, 1101-1108.

[24] McCallum, B. (2009b), Rejoinder to Cochrane. Journal of Monetary Economics 56, 1114-1115.

[25] McCallum, B. (2012), Indeterminacy, learnability, plausibility and the role of money in new Keynesian models, NBER Working Paper 18125.

[26] McGough, B., Meng, Q. and Xue, J. (2013), Expectational stability of sunspot equilibria in nonconvex economies. Journal of Economic Dynamics and Control 37, 1126-1141.

[27] Preston, B. (2005), Learning about monetary policy rules when long-horizon expectations matter, International Journal of Central Banking 1, 81-126.

[28] Sargent, T.J. (2008), Evolution and intelligent design, American Economic Review 98, 5-37.