Observability and Equilibrium Selection

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

Bruce McGough
University of Oregon

August 28, 2015

Abstract

We explore the connection between shock observability and equilibrium. When aggregate shocks are unobserved the rational expectations solutions remain unchanged if contemporaneous aggregate outcomes are observable, but their stability under adaptive learning must be reconsidered. We obtain learning stability conditions and show that, provided there is not large positive expectational feedback, the minimum state variable solution is robustly stable under learning, while the nonfundamental solution is never robustly stable. Using agent-level micro-founded settings we apply our results to overlapping generations and New Keynesian models, and we address the uniqueness concerns raised in Cochrane (2011).

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

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

1 Introduction

We explore the connection between shock observability and equilibrium selection under learning. Under rational expectations (RE) the assumption of shock observability does not affect equilibrium outcomes if current aggregates are in the information set. However, 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.

Within the context of a simple model, we establish that determinacy implies the existence of a unique robustly stable rational expectations equilibrium. Our results also cover the indeterminate case. As an important application of our results, we examine the concerns raised in Cochrane (2009, 2011) about the New Keynesian model.

A rational expectations equilibrium (REE) may be defined as a sequence of conditional distributions for a model's endogenous variables that are consistent with the decisions of agents who themselves are making forecasts optimally against these distributions. Importantly, for an REE to be well-defined, the information sets available to agents when forecasts are made must be precisely specified. Most often, the information sets are implicitly specified during the construction of the corresponding model.

Even while being sharply criticized both theoretically and empirically, the assumption of rational expectations remains standard among researchers and policy makers, and serves as the primary benchmark against which any natural alternative is compared. The reasons for the profession's love-affair with rationality are (at least) two-fold: first, from a strictly practical point of view, rational expectations models are straightforward to analyze; second, and much more fundamentally, RE, by definition, uniquely identifies an expectations formation mechanism that the associated agents cannot improve upon – there are no twenty dollar bills left on the roads of RE models.

Among the theoretical criticisms levied at the RE assumption, three stand out: the unrealistic structural knowledge apparently required by agents; the need for a story for how agents coordinate on an REE; and the potential for equilibrium multiplicity. The first two criticisms can be described together. When economists compute an REE we do so with knowledge of the full structure of the economy, including all the relevant structural parameters. It does not seem plausible that agents would have the needed knowledge of economy-wide parameters, given especially that in practice economists need to estimate them. In addition, a crucial feature of macroeconomic models is that outcomes depend on expectations, introducing a self-referential aspect to the economy: it is rational to have RE, i.e. the expectations corresponding to an REE, only if other agents have RE. Through what mechanism would agents come to coordinate on having RE?

Finally there is the issue of multiplicity of RE solutions. Even simple linear RE models typically have multiple RE solutions. Which solution should be chosen? In many cases (when "determinacy" holds) there is a unique non-explosive RE solution, and in this case it is common practice to pick this 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 in which there are multiple non-explosive RE solutions. The latter can take various forms, including models with a unique but "indeterminate" steady state (with multiple stationary RE solutions in any neighborhood of the steady state) and models with multiple RE steady states.

The literature on adaptive learning arose in part to mitigate these concerns. Under this approach agents are assumed to form expectations using forecasting models, which they update over time in response to observed data. The now well-developed theory of adaptive learning 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 rational expectations. In this case we say that the associated equilibrium is stable under adaptive learning. It has been found, in many cases, that stability under adaptive learning can be used to justify reliance on RE: see Marcet and Sargent (1989) and Evans and Honkapohja (2001) for a careful treatment.

In addition, in settings with multiple REE, adaptive learning has played an effective role as an equilibrium selection device: reducing consideration to REE that are locally stable under learning dynamics can either select a unique equilibrium or at least restrict attention to a subset.[1] Even in the cases in which multiplicity remains, adaptive learning resolves the multiplicity problem in the sense that it provides a full description of the path of the economy that is generated when agents use a given forecasting model, and updating procedure, and when initial beliefs are specified.

Under the adaptive learning approach a stand must still be take on what information is available to forecasting agents. The common, though by no means ubiquitous, assumption is that agents are able to condition time t forecasts on t-dated variables, including any t-dated exogenous shocks. While in some environments this assumption may be quite natural, in others it is quite difficult to defend. Should aggregate productivity shocks be taken as observable? Certainly one cannot look up the current value in a newspaper. Perhaps the shocks can be "backed-out" using, e.g. the Solow residual, but this requires estimation, and so at best, we might take a proxy as observable. And what about taste shocks? Are these observable? Even shocks to someone else's taste? These issues have been raised recently in policy-related literature: Cochrane (2011) has argued that the New Keynesian model's exogenous shocks, specifically the monetary policy shocks, are most naturally taken as unobservable, and that this unobservability has important consequences: more on this later.

Of course it is not necessary to abandon RE even if one imagines that shocks, at least contemporaneous shocks, should not be assumed observable: the information sets available to agents can simply be specified to exclude certain variables and, in this way, RE can still be embraced. Lucas (1973) did precisely this in his famous

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

islands model, and Mankiw and Reis (2002) have exploited the same idea in their sticky information DSGE environment. In this paper, we explore the connection between observability and equilibrium selection. We take as our point of departure the possibility that contemporaneous exogenous shocks are not observable to agents.

In Section 2 we consider a simple forward-looking model with an exogenous serially correlated shock, and formulate the REE in two cases: when the shock is observable and when it is not. In setting up the framework we look ahead to the subsequent sections on stability under adaptive learning, and define an REE as requiring agent beliefs to satisfy the property that the probability distribution of the resulting data generating process coincide with the beliefs.[2] Starting with the case with observable shocks, we identify two REE of interest: the fundamental, or "minimum state variable" (MSV) solution, which is always stationary; and a non-fundamental (NF) solution, which may or may not also be stationary. We then show that when shocks are not observable there are RE beliefs that reproduce these solutions. The forecast rules that generate the REE use lags of the observable endogenous variable in effect to uncover the information in the unobservable exogenous variables needed to make optimal forecasts.

We then, in Section 3, turn to stability under learning. In a model with observable shocks, McCallum (2009a) showed that determinacy implies that only the fundamental MSV solution is stable under learning. However, Cochrane (2009, 2011) argued that McCallum's stability results hinged on observability of these shocks. In Section 3 we conduct a careful study of stability under learning of the two solutions when the exogenous shocks are unobservable. Since it will not be obvious to agents how many lags of the observable endogenous variable to use in the specification of their forecasting model, our criterion for stability under learning, which we call "robust stability," is that the REE be locally stable under learning even if the forecast rule contains more than the minimum number of lags needed to represent the REE. The central result of our paper is that 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.

Section 4 studies two applications. Both applications are developed from microfounded agent-level models. In both examples a key step is to obtain what is called the "temporary equilibrium map," which gives the time t market outcomes, i.e. the market-clearing prices and quantities, that result from the supply and demand schedules of agents conditional on their expectations of relevant variables.[3] In each case the temporary equilibrium map, when linearized, matches the framework developed Sections 2 and 3.

[2] In the language of Adam and Marcet (2011) the beliefs are externally as well as internally rational.

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

The first application of Section 4 is an overlapping generations model with production and money. This leads to a temporary equilibrium map in which the goods price of money depends on its expected price in the following period. This model can be either determinate (in which case the MSV solution is the unique stationary solution) or indeterminate (in which case the NF solution is also stationary). We show that in either case the MSV solution is robustly stable under learning and the NF solution is not.

The second application is a New Keynesian (NK) model. We use the flexible-price NK model employed by both McCallum (2009a) and Cochrane (2009) in their interchange. Again, we build up the model from the agent level to obtain the system of equations that define the temporary equilibrium, and then we discuss how the temporary-equilibrium variables are affected by a change in expected inflation. The linearized model again can be put in the form analyzed in the earlier sections, and provided that 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.

To summarize: adaptive learning provides a natural mechanism by which agents can learn to coordinate on an REE. In the two applications we study, it also acts as a selection criterion even if the exogenous shocks are unobservable. In particular, in the NK model adaptive learning selects the REE solution typically adopted by practitioners.

2 Observability and rational expectations

Is it natural to assume exogenous shocks are observable? Certainly it is in some cases. For example, it is common to take the exogenous components of government spending and taxes as observable to forecast-forming agents. One the other hand, assuming agents condition their expectations on technology shocks (which economists themselves can only estimate) or taste shocks (which are only revealed ex-post via choices) is more difficult to defend. The examination of models under different informational restrictions has a long history. An early example is provided by Lucas (1973): he assumed aggregate price levels were unobserved, and he leveraged this friction to impart real effects of surprise money on output; these same types of informational frictions underlie the modern literature on rational inattention: see Mankiw and Reis (2002) and Sims (2003).

The specific assumption that certain exogenous shocks are unobserved has also been explored and exploited in a variety of models. Within the Real Business Cycle framework, it is common to assume certain types of productivity shocks are unobserved: see King and Rebelo (1999) for examples and further references. Similar examinations have been conducted in the New Keynesian model: Woodford (2003) considers the possibility that expectations are formed before certain shocks are realized; and more recently, and again in a New Keynesian model, Levine et al (2012) find that the assumption that shocks are not observed by agents leads to improved empirical performance.

Our interest here is in the relationships among observability, equilibrium multiplicity, and equilibrium selection. The most closely related works are Cochrane (2009) and McCallum (2009a,b). In his papers, Cochrane considers a simple New Keynesian model, and argues that the monetary policy shock (which is captured by an innovation associated to an instrument rule) should not be taken as observable; and he finds that, in this case, there are multiple learnable equilibria even when the model is determinate. Cochrane's work motivates some of our work here, and we will have more to say on his results and related issues in Section 4.2.

2.1 The model

To explore the relationship between observability and expectations formation within an RE framework, we adopt a simple, linear model as our laboratory, written in a manner consistent with the literature on rational expectations:

$$\begin{aligned} y_t &= \beta E_t y_{t+1} + v_t && (1) \\ v_t &= \rho v_{t-1} + \varepsilon_t, \text{ where } 0 < \rho < 1, && (2) \\ \mathcal{B} &= \text{boundary conditions} \end{aligned}$$

Here v_t captures a positively autocorrelated stationary exogenous process, with ε_t zero mean, *iid* and having bounded support.[4] 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.[5] Finally, the symbol $\mathcal{B}$ represents the (possibly empty) set of all additional boundary conditions needed to characterize the collection of equilibrium outcomes: for example, it may be required that candidate processes y_t are asymptotically stationary or uniformly bounded almost everywhere.

In this section we embrace the rational expectations hypothesis, which may be characterized as follows: agents are assumed to form forecasts optimally given their available information and beliefs, and in (a rational expectations) equilibrium, agents have no incentive to modify their beliefs. The symbol E_t is typically taken to represent

[4] We focus on the case of positively serially correlated shocks as this is the empirically realistic case. Negative serial correlation of v_t would not make the NF solution robustly stable under learnng.

[5] 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.

the conditional expectations operator, where the conditioning is taken against the σ-algebra $\mathcal{F}_t$ generated by the collection of all random variables assumed observable at the time t, that is, when expectations are formed. In general, agents could also condition their forecasts on extrinsic shocks (i.e. "sunspots"), in which case these variables would necessarily be taken as observable and thus measurable against the σ-algebra $\mathcal{F}_t$. Because our focus here is on the implications of not observing exogenous shocks, in this paper we exclude expectations conditioned on sunspots.

If the values of both the endogenous and exogenous variables are known to agents then our interpretation of E_t as a conditional expectations operator raises no issues; however, a subtlety arises when the exogenous shocks are unobserved: in this case, v_t will be measurable with respect to $\mathcal{F}_t$, but realizations of v_t cannot be explicitly used by agents to form forecasts. To deal with this subtlety we draw inspiration from Adam and Marcet (2009) and develop a precise specification of rational expectations, and the associated equilibrium concept, in two stages: first, we require that agents be *internally rational* in that they form expectations optimally given their beliefs; and second, in equilibrium we require that the agents be *externally rational*, i.e. their beliefs coincide with the distribution implied by the true data-generating process (DGP). The related but distinct behavioral assumptions of internal and external rationality do not introduce a new solution concept – the associated equilibrium is an REE, and in any REE agents are both internally and externally rational; instead, as will be seen in the following subsections, these behavioral assumptions allow for a careful and rigorous approach to the incorporation of structural assumptions on agents' forecasting models.

2.2 RE with observable shocks

Let Y and V be copies of $\mathbb{R}^\infty$, that is, vector spaces of real sequences. Assume that agents are identical, and hold beliefs characterized by a probability distribution $\mathcal{P}$ over $Y \times V$, that is, over possible realizations of the pair of stochastic processes $(\{y_t\}, \{v_t\})_{t\in\mathbb{N}}$. Letting y^t and v^t be the respective history vectors, the beliefs $\mathcal{P}$ determine the perceived distribution of y_{t+1} conditional on the histories y^t and v^t. We say that agents are internally rational if they form expectations against these perceived conditional distributions: $E_t y_{t+1} = E^{\mathcal{P}}(y_{t+1}|y^t, v^t)$. The superscript $\mathcal{P}$ tracks the underlying joint distribution being used to form the expectations; similar notation is used below.

Incorporating internally rational agents with beliefs $\mathcal{P}$ into model (1) yields the data-generating process

$$y_t = \beta E^{\mathcal{P}}\left(y_{t+1}|y^t, v^t\right) + v_t. \tag{3}$$

This DGP determines a realized probability distribution $T(\mathcal{P})$ over $Y \times V$ that depends on the agents' beliefs $\mathcal{P}$, as is indicated by the notation used. We say that internally rational agents are also externally rational provided that the realized distribution coincides with their beliefs, i.e. $T(\mathcal{P}) = \mathcal{P}$. If agents are externally rational, and if the corresponding process y_t, as determined by (3), satisfies the boundary conditions $\mathcal{B}$, then y_t is a rational expectations equilibrium in the usual sense of the term. Formally,

Definition 1 *A* rational expectations equilibrium *of the model (1), under the assumption that v^t is observable, is a probability distribution $\mathcal{P}$ over $Y \times V$ and a stochastic process y_t satisfying equations (3) and $T(\mathcal{P}) = \mathcal{P}$, and the conditions $\mathcal{B}$.*

We again emphasize that no new equilibrium concept has been introduced: if y_t is an REE of (1) in the usual sense then choosing beliefs $\mathcal{P}$ to correspond to the associated distribution over $Y \times V$ shows that it is is an REE in the sense of Definition 1; and if y_t is the stochastic process associated to an REE in the sense of Definition 1 then $T(\mathcal{P}) = \mathcal{P}$ implies that agents are forecasting optimally along the equilibrium path; thus they have no incentive to modify their beliefs and y_t is an REE in the usual sense. The distinction between Definition 1 and the usual definition of an REE is the emphasis placed on agents' beliefs.

To operationalize Definition 1, and thus use it to compute equilibria, it is helpful to stress how certain beliefs structures $\mathcal{P}$ may be constructed from, and characterized by, forecasting models. Suppose agents believe that the process v_t is given by (2) and that y_t follows

$$y_t = a + \sum_{i=1}^{N} \phi_i y_{t-i} + \sum_{j=0}^{M} \lambda_j v_{t-j}. \tag{4}$$

In general we will refer to a forecasting model like this as a Perceived Law of Motion (PLM).[6] Together with initial conditions, (4) and (2) characterize a distribution over $Y \times V$. As terminology we say that the PLM generates the distribution $\mathcal{P}$. The implied forecasts are given by

$$E^{\mathcal{P}}\left(y_{t+1}|y^t, v^t\right) = a + \sum_{i=1}^{N} \phi_i y_{t+1-i} + \sum_{j=1}^{M} \lambda_j v_{t+1-j} + \rho\lambda_0 v_t,$$

where agents are assumed to know ρ and use (2) to forecast v_{t+1}. We emphasize that agents are assumed to use y_t to forecast y_{t+1}.[7] These expectations can be joined with equation (3) to determine the DGP, which is also referred to as the Actual Law of Motion (ALM). Note that the PLM includes an intercept, so that under adaptive

[6] Agents engaged in real-time least-squares learning would also include a random error term.

[7] An alternative assumption sometimes considered is that the exogenous shock is observable but the contemporaneous aggregate endogenous variable is not observable. We do not consider this case since our focus is on the implications of not observing exogenous shocks.

learning agents will in effect need to learn the steady state as well as the other coefficients.

If $\mathcal{B}$ is empty it is straightforward to show that there are necessarily many equilibria. We emphasize two here: the minimal state variable (MSV) solution, sometimes also called the fundamentals solution, and the particular non-fundamental (NF) solution studied by McCallum (2009a) and Cochrane (2009).[8] When the model is determinate, the MSV solution is the unique stationary solution and all non-fundamental solutions are explosive and are often referred to as bubbles.

The MSV solution arises from beliefs generated by a forecasting model of the form $y_t = \lambda_0 v_t$. When $\lambda_0 = (1-\beta\rho)^{-1}$, forecasts based on $\mathcal{P}$ and $T(\mathcal{P})$ coincide. The NF solution that we consider corresponds to a forecasting model that also includes a perceived dependence on lags:

$$y_t = \phi_1 y_{t-1} + \lambda_0 v_t.$$

Here, when $\phi_1 = \beta^{-1}$ and $\lambda_0 = -(\beta\rho)^{-1}$ forecasts based on $\mathcal{P}$ and $T(\mathcal{P})$ are coincident. For reference, we name and record the associate equilibrium processes here:

$$\text{MSV:} \qquad y_t = (1-\beta\rho)^{-1} v_t \tag{5}$$

$$\text{NF:} \qquad y_t = \beta^{-1} y_{t-1} - \frac{1}{\beta\rho} v_t. \tag{6}$$

The MSV solution is the REE exhibiting dependence on a minimal set of regressors.[9] The NF solution, which is not minimal in the sense that it also depends on y_{t-1} (and is self-fulfilling in the sense that the lag is present if and only if agents perceive its presence), is explosive almost surely when $\beta \in (-1, 1)$, i.e. when the model is determinate.

Most macroeconomic models of the form (1) include a non-trivial set of boundary conditions $\mathcal{B}$. These conditions may arise from physical constraints (e.g. consumption cannot be negative), institutional constraints (e.g. borrowing cannot be greater than a certain proportion of collateral value), or additional optimality conditions (e.g. TVC). The presence of boundary conditions may or may not preclude the existence of multiple equilibria. If, in our example above, $\mathcal{B}$ specifies only that solutions not grow at a rate faster than $\beta^{-1} + \delta$ for some $\delta > 0$ then both the MSV and NF solutions are REE even in the determinate case. If, on the other hand, $\mathcal{B}$ includes the requirement that the sequence $\{y_t\}$ be uniformly bounded almost surely, then, in the determinate case, the NF solution is not an REE.

[8] See McCallum (2009a), equation (9) on p. 1103 and Cochrane (2009), equation (11) on p. 1112. There are other NF solutions consistent with (4): see Appendix B for details.

[9] Using the method of undetermined coefficients, it is straightforward to show, for this model, that the MSV solution is unique. McCallum (1983) shows how to define a unique MSV solution in more general models that include lags.

This latter condition, that REE be uniformly bounded almost surely, is commonly imposed by modelers: it is seen as capturing the need for the solution to satisfy agents' TVCs, and possibly for the solution to respect that the model under examination is really a first-order approximation to a nonlinear system. Because of its prominence in the literature, we label and examine this condition more closely, and within the context of the related notion, *determinacy.*

Definition 2 *A stochastic process* $\{y_t\}$ *satisfies the* uniform boundedness condition (UBC) *if there exists* $M > 0$ *so that* $\|y_t\|_\infty < M$ *for all* $t \geq 0$*.*[10]

The UBC is a standard, but not ubiquitous boundary condition; however, we know of no common boundary conditions that are more restrictive. For this reason, we will assume that if a potential solution satisfies the UBC then it satisfies all conditions $\mathcal{B}$; in particular, the MSV solution always satisfies the conditions $\mathcal{B}$.

The closely related concept of determinacy is defined as follows:

Definition 3 *The model (1) is* determinate *if there is a unique REE satisfying the UBC.*[11]

We note that if $|\beta| < 1$ then the NF solution does not satisfy the UBC. In fact, we have the following well-known result:

Proposition 1 *The model (1) is determinate if and only of* $|\beta| < 1$.

All proofs are in the Appendix.

Determinacy is an appealing model characteristic: it is straightforward to assess; and it implies, in a natural sense, local uniqueness of the REE. In fact, determinacy is often taken as equivalent to uniqueness of REE in macroeconomic models. However, as our discussion in this Section emphasizes, if $\mathcal{B}$ does not include the UBC (or some closely related condition) then determinacy does not necessarily imply that the model pins down the equilibrium. Cochrane (2011) makes this point forcefully, where he argues that determinacy in the New Keynesian model does not imply the existence of a unique REE; in fact, according to Cochrane, certain bubble solutions satisfy the boundary conditions imposed by the model. We will return to this point later.

[10] Here, $\|\cdot\|_\infty$ refers to the essential supremum norm, which, in the parlance of probability theory, may be defined as follows: if x is a real-valued random variable then

$$\|x\|_\infty = \inf\{M \in \mathbb{R} : Prob(|x| > M) = 0\}.$$

[11] If the model is non-linear, determinacy is a local notion.

2.3 RE with unobservable shocks

As emphasized by Cochrane (2009, 2011), it may not be natural or reasonable to assume that a given model's exogenous shocks are observable. To model shocks as fundamentally unobservable, we restrict agents beliefs $\mathcal{P}$ to distributions only over Y (instead of $Y \times V$), and allow them to form forecasts conditional only on current and past values of y: $E_t y_{t+1} = E^{\mathcal{P}}(y_{t+1}|y^t)$. The DGP implied by beliefs $\mathcal{P}$ is given by

$$y_t = \beta E^{\mathcal{P}}\left(y_{t+1}|y^t\right) + v_t, \tag{7}$$

which determines a realized probability distribution $T(\mathcal{P})$ over Y. This leads to the following analogue to Definition 1:

Definition 4 *A* rational expectations equilibrium *of the model (1), under the assumption that v^t is not observable, is a probability distribution $\mathcal{P}$ over Y and a stochastic process y_t satisfying equations (7) and $T(\mathcal{P}) = \mathcal{P}$, and the conditions $\mathcal{B}$.*

We note that an equilibrium in the sense of Definition 4 is an equilibrium in the sense of Definition 1: if $\mathcal{P}$ is an equilibrium distribution over Y generating the equilibrium process y_t then the joint distribution $\mathcal{P}'$ over $Y \times V$ implied by (7) defines an equilibrium in the sense of Definition 1. In particular, if the model is determinate then there will be at most one equilibrium satisfying the UBC regardless of the choice of definitions.

The MSV and NF solutions are equilibria in the sense of Definition 4. This can be seen most easily by using (2) to represent these solutions as follows

$$\text{MSV:} \qquad y_t = \rho y_{t-1} + (1 - \beta\rho)^{-1}\varepsilon_t \tag{8}$$

$$\text{NF:} \qquad y_t = \left(\beta^{-1} + \rho\right) y_{t-1} - \beta^{-1}\rho y_{t-2} - \frac{1}{\beta\rho}\varepsilon_t. \tag{9}$$

Motivated by these representations, we now consider belief structures that will support these solutions as REE in the sense of Definition 4.

As in the previous section, specific beliefs $\mathcal{P}$ are predicated upon forecasting models. It will be useful to consider PLMs allowing for a finite number of lags of y, and so we modify the PLM (4) accordingly:

$$y_t = a + \sum_{n=1}^{N} \phi_n y_{t-n} + \xi_t, \tag{10}$$

where ξ_t represents the agent's 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 $\psi = (a, \phi)$, where $\phi = (\phi_1, \ldots, \phi_N)$.

Given their beliefs, agents form expectations optimally:

$$E_t^{\mathcal{P}}\left(y_{t+1}|y^t\right) = a + \sum_{n=1}^{N} \phi_n y_{t+1-n}.$$

Again, we emphasize that 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 (2) and letting $\gamma(\phi_1) = (1-\beta\phi_1)^{-1}$, we have the ALM

$$\begin{aligned}
y_t &= \gamma(\phi_1)\beta a(1-\rho) + \rho y_{t-1} + \gamma(\phi_1)\varepsilon_t, \text{ for } N = 1 \\
y_t &= \gamma(\phi_1)\beta a(1-\rho) + (\rho + \gamma(\phi_1)\beta\phi_2) y_{t-1} - (\gamma(\phi_1)\beta\rho\phi_2)\, y_{t-2} + \gamma(\phi_1)\varepsilon_t, \text{ for } N = 2 \\
y_t &= \gamma(\phi_1)\beta a(1-\rho) + (\rho + \gamma(\phi_1)\beta\phi_2) y_{t-1} + \gamma(\phi_1)\beta \sum_{n=2}^{N-1} (\phi_{n+1} - \rho\phi_n) y_{t-n} \\
&\quad - (\gamma(\phi_1)\beta\rho\phi_N)\, y_{t-N} + \gamma(\phi_1)\varepsilon_t, \text{ for } N \geq 3.
\end{aligned}$$

The functional forms of the forecasting model and the data generating process exhibit the same linear dependency. This allows us to identify a map

$$T_N : \mathbb{R} \oplus \mathbb{R}^N \to \mathbb{R} \oplus \mathbb{R}^N$$

between the respective coefficients, where the subscript N on the map tracks the number of lags of y in the PLM. Just as the PLM generates the distribution $\mathcal{P}$, so the ALM generates the distribution $T(\mathcal{P})$. It follows that a fixed point of the map T_N identifies a fixed point of T.

The following notation will be useful: set

$$\psi^{MSV} = (0, \rho) \text{ and } \psi^{NF} = (0, \beta^{-1} + \rho, -\beta^{-1}\rho),$$

corresponding to equations (8) and (9). Define

$$\begin{aligned}
\psi_N^{MSV} &= (\psi^{MSV}, \underbrace{0, \ldots, 0}_{N-1 \text{ terms}}) \text{ for } N \geq 1, \\
\psi_N^{NF} &= (\psi^{NF}, \underbrace{0, \ldots, 0}_{N-2 \text{ terms}}) \text{ for } N \geq 2.
\end{aligned}$$

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

Proposition 2 *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}.*

This proposition shows that there are precisely two equilibrium processes y_t consistent with forecasting models of the form (10), and that they correspond to the MSV and NF solutions. In particular, both the MSV solution and the NF solution may be realized as REE in the sense of Definition 4.

3 Observability and equilibrium selection

In models of the form (1) Cochrane (2011) and others have noted that the set of conditions $\mathcal{B}$ which organically emerge from a micro-founded model may not be sufficiently restrictive to preclude the relevance of the NF solution: determinacy may not determine the equilibrium path. In such cases, i.e. when equilibrium multiplicity prevails, adaptive learning offers a natural equilibrium selection mechanism. McCallum (2007, 2009a) has shown, using E-stability analysis along the lines of Section 3.1, below, that if the exogenous shocks are taken as observable then only the MSV solution is learnable: he concludes that the NF solution is of little interest and that determinacy is the appropriate condition to ensure that there is one relevant equilibrium.[12] Cochrane subsequently argued that, at least within the context of a simple New Keynesian model, it is unrealistic to assume the exogenous shocks are observable, and he suggests that if they are taken as unobserved then, in fact, the NF solution will be learnable and that the MSV solution is not learnable. In the current section we adopt this position on observability and assess whether the MSV solution and the NF solution are learnable.

3.1 Adaptive learning and E-stability

Our discussion and development will be compact; for details see Marcet and Sargent (1989) and Evans and Honkapohja (2001). We focus here on the case of unobserved shocks. Under adaptive learning agents use available data to estimate a forecasting model of the form (10). Thus at time t agents have parameter estimates $\psi_t = (a_t, \phi_t)$, obtained using data through time $t-1$, and under adaptive learning $E_t^{\mathcal{P}}(y_{t+1}|y^t)$ is replaced by forecasts[13]

$$\hat{E}_t y_{t+1} = a_t + \sum_{n=1}^{N} \phi_{t,n} y_{t+1-n}.$$

The implied DGP is then given by

$$y_t = \beta \hat{E}_t y_{t+1} + v_t.$$

The additional data point y_t at time t is then used to update the forecast model coefficients to ψ_{t+1}, which can be done efficiently using recursive least squares. The updated coefficients are then used in the forecast rule in period $t+1$ and the process continues. This model of expectation formation adheres to the "cognitive consistency

[12] In general, when the model is determinate, when shocks are observable and when the model has no endogenous lags, the MSV solution is the unique E-stable solution. See also Evans (1989).

[13] Adam and Marcet (2009) show that in certain models adaptive learning is consistent with internal rationality.

principle" in which the level of bounded rationality of agents is in line with that of economists themselves.[14]

If the resulting forecasts and the implied decisions induce model dynamics that converge in an appropriate sense to a particular REE, that REE is said to be stable under learning. Many authors, including McCallum (2007, 2009a), have argued forcefully and effectively that stability under learning is a natural and reasonable equilibrium selection mechanism for macroeconomic models: if a particular REE is stable under learning then it may be viewed as a plausible equilibrium outcome for the model under examination; and if, among multiple equilibria, there is a unique stable REE then the model may be considered to produce a precise outcome.

The examination of stability under learning is greatly facilitated by the E-stability principle as developed, for example, in Evans and Honkapohja (2001), and which we now apply in the current setting. We imagine that agents use a forecasting model of the form (10), for some $N \geq 1$. Summarizing the beliefs coefficients by $\psi = (a, \phi)$, we may consider the following system of differential equations:

$$\frac{d\psi}{d\tau} = T_N(\psi) - \psi, \tag{11}$$

where τ is a notional time variable (which can, however, be linked to calendar time). A rest point of this differential system coincides with a fixed point of the map T_N; thus both the MSV solution and the NF solution may be represented as rest points of (11).

An REE corresponding to a fixed point ψ of T_N is said to be *E-stable* if ψ a Lyapunov-stable rest point of (11). According to the *E-stability Principle*, an E-stable REE is (locally) stable under adaptive learning provided agents use recursive least-squares or closely related updating algorithms. While the connection between the system of differential equations (11) and the convergence of recursive learning algorithms is quite deep, the intuition is straight-forward: The right-hand-side of (11), $T_N(\psi) - \psi$, may be thought of as "truth" minus "perception," that is, a type of forecast error. The system (11) moves beliefs in the direction of this forecast error, and this movement leads to convergence to the (fixed-point associated to the) REE provided that the REE is E-stable. Since learning algorithms behave similarly in that they adjust parameter estimates in the direction indicated by the forecast error, their convergence is also, then, implied.

3.2 Robust stability and equilibrium selection

It is not a priori obvious how many lags agents should include in their forecasting model, and because of this, our equilibrium selection mechanism requires modifica-

[14] See Evans and Honkpohja (2009).

tion. Recall that the subscript N in ψ_N^{MSV} and ψ_N^{NF} denotes the number of lags of y in the agent's PLM. We have the following definition:

Definition 5 *The MSV solution is robustly stable under learning provided ψ_N^{MSV} is a Lyapunov-stable rest point of (11) for all $N \geq 1$. The NF solution is robustly stable under learning provided ψ_N^{NF} is a Lyapunov-stable rest point of (11) 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 for its specification. Since it is not obvious how many lags agents should/would include in their forecasting models, we think of robust stability as a natural selection mechanism for this and related models.

We have the following theorem, which is the main result of the paper:

Theorem 1 *If the exogenous shocks are not observable in the model (1) then*

1. *the MSV solution is robustly stable under learning when $\beta < 1$ and not robustly stable under learning if $\beta > 1$;*
2. *the NF solution is never robustly stable under learning.*

The proof is in the Appendix, and its subtlety warrants comment. Assume, for this discussion, that $\mathcal{B}$ is empty. Then both the MSV solution and the NF solution are REE. It is reasonably straightforward to show that the MSV solution is robustly stable. Showing the NF solution is not robustly stable requires more work: for $N = 2$ the NF solution does correspond to a Lyapunov-stable fixed point; however, for $N \geq 3$ the differential system (11) has an eigenvalue with zero real part, and a center-manifold reduction argument yields that the rest point of (11) corresponding to the NF solution is not Lyapunov stable. Thus in fact 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 $N \geq 3$. Furthermore, we note that when $\beta > 1$ no fixed point of T_N is Lyapunov stable; that is, when $\beta > 1$ no solution is stable under learning.

Theorem 1 applies whether the model is determinate or indeterminate. In either case the MSV solution is stationary and robustly stable under learning. In the indeterminate case the NF solution is also stationary while in the determinate case it is an explosive 'bubble.' However in both cases the NF solution fails to be robustly stable under learning.

McCallum (2007) showed that determinacy implies E-stability of the MSV solution, and McCallum (2009a) established that the NF solution is not E-stable. A number of authors including Cochrane have questioned whether it is natural to assume that exogenous shocks are observable; furthermore, Cochrane (2011) argues that 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 1 dispels this notion, and extends McCallum's result to unobservable shocks.[15]

4 Applications

The generality of the simple, forward-looking model (1) allows for a variety of applications. In this subsection, we consider two: the overlapping generations model, as developed in Section 4.1, provides one natural example, in which we emphasize the indeterminate case; and a simplified version of the benchmark New Keynesian model, as examined by Cochrane (2011), in which the determinate case is examined.

4.1 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

$$\max_{c_{t+1}(\omega_t),\ n_t(\omega_t),\ M_t(\omega_t)} \quad \hat{E}(\omega_t)\left(u(c_{t+1}(\omega_t)) - \nu(n_t(\omega_t))\right) \tag{12}$$

$$\text{subject to} \quad z_t(\omega_t)n_t(\omega_t) = q_t M_t(\omega_t) \;\text{ and }\; c_{t+1}(\omega_t) = q_{t+1}M_t(\omega_t)$$

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

[15] The assumption that contemporaneous values of the endogenous variables are used when forecasts are made is central here. To assume otherwise, i.e. that no contemporaneous variables are observable, raises a number of fundamental issues regarding the appropriate notion of equilibrium. This is a topic of current research.

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.

The first order condition is given by

$$\nu'(n_t(\omega_t)) = \hat{E}(\omega_t)\left(z_t(\omega_t)\frac{q_{t+1}}{q_t}u'(c_{t+1}(\omega_t))\right), \tag{13}$$

and to make our model particularly tractable, we assume that $\nu' = 1$ and

$$u(c) = \frac{1}{1-\sigma}\left(c^{1-\sigma} - 1\right).$$

With simplification, we obtain agent ω_t's decision rules:

$$\begin{aligned} n_t(\omega_t) &= \left(\left(\frac{z_t(\omega_t)}{q_t}\right)^{1-\sigma}\hat{E}(\omega_t)\left(q_{t+1}^{1-\sigma}\right)\right)^{\frac{1}{\sigma}} \\ M_t(\omega_t) &= \left(\frac{z_t(\omega_t)}{q_t}\hat{E}(\omega_t)\left(q_{t+1}^{1-\sigma}\right)\right)^{\frac{1}{\sigma}}; \end{aligned}$$

and we note that, as is natural, 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, we obtain the market-clearing condition

$$\int_\Omega M_t(\omega_t)d\omega_t = 1,$$

which yields

$$q_t = \left(\int_\Omega \left(z_t(\omega_t)\hat{E}(\omega_t)\left(q_{t+1}^{1-\sigma}\right)\right)^{\frac{1}{\sigma}}d\omega_t\right)^{\sigma}. \tag{14}$$

Equation (14) characterizes the equilibrium price path.

As written, the system (14) is quite complicated, involving, as it does, aggregation, expectations and non-linearity. Also, the model is not yet closed: the details of expectations formation have not been specified. The standard protocol is to impose rational expectations. This assumption eliminates heterogeneity, thus simplifying aggregation, and the equilibrium price path may then be approximated by linearizing the dynamic system around a non-stochastic steady state and applying standard techniques. Here, to emphasize the particulars of observability, we first provide a useful specification for the productivity shocks; then we linearize the model and contemplate rationality.

We assume agent ω_t's productivity shock includes both an idiosyncratic and an aggregate component. Specifically

$$\begin{aligned} \log(z_t(\omega_t)) &= \log(z_t) + \log(\zeta_t(\omega_t)) \\ \log(z_t) &= \rho\log(z_{t-1}) + \varepsilon_t, \end{aligned}$$

where the $\log(\zeta_t(\omega_t))$ are *iid* mean zero and independent across agents, ε_t is *iid* zero mean with small support and the $\log(\zeta_t(\omega_t))$ and ε_t are independent processes.[16] An implication of this assumed structure is that

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

In particular, note that for every agent ω_t the contribution of $\zeta_t(\omega_t)$ to q_t is negligible. It follows that $z_t(\omega_t)$ can be useful for forecasting q_{t+1} only through is 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.

Under RE the equilibrium price process can be represented as depending explicitly on the z_t process. Indeed, an REE is characterized by

$$q_t = \kappa z_t E_t q_{t+1}^{1-\sigma}, \text{ where } \kappa^{\frac{1}{\sigma}} = \int_\Omega \zeta_t(\omega_t)^{\frac{1}{\sigma}} d\omega_t.$$

However, it is neither natural nor necessary to assume agents observe z_t when making rational forecasts. What the agent needs to predict is the future price level q_{t+1} and, along the lines of the Section 2.3 discussion, 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.

Returning to the analysis of the model, all agents are assumed to 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. With this assumption, 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 (14) around this steady state and simplify to obtain

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

Noting that, by independence, $\int_\Omega \log z_t(\omega_t) d\omega_t = \log z_t$, we obtain

$$\log q_t = (1-\sigma)\hat{E}_t \log q_{t+1} + \log z_t, \tag{16}$$

which reduces to our model (1) with $y_t = \log q_t$ and $v_t = \log z_t$. We remark that the model (16) is indeterminate when $\sigma > 2$.

[16]The assumption that $\log(\zeta_t(\omega_t))$ are *iid* is stronger than needed; however, it is essential that they are mean zero and independent across agents.

The analysis of Section 2 implies that the model (16) has both the MSV solution and the NF solution as stationary equilibria when $\sigma > 2$. However, Theorem 1 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 converge either to infinity or zero almost surely. Whether the explosive path represents a genuine equilibrium depends on interpretation of model details, but in any case the NF solution is not robustly stable under learning.[17]

4.2 Equilibrium selection in the New Keynesian model

We investigate the application of our results to a New Keynesian (NK) environment. The standard version of this is the three-equation linearized NK model

$$\begin{aligned} x_t &= E_t x_{t+1} - \sigma^{-1}(R_t - E_t \pi_{t+1}) \\ \pi_t &= \delta E_t \pi_{t+1} + \varphi x_t \\ R_t &= \alpha_\pi \pi_t + \alpha_x x_t + v_t, \end{aligned}$$

where x_t is the output gap, R_t is the nominal interest rate, and π_t is the inflation rate, with all variables written in proportional deviation from mean form. The first equation is the NK "IS curve," based on the consumption Euler equation. The second equation is the NK "Phillips curve," where here $0 < \delta < 1$ is the household's discount factor. The third equation is the standard Taylor rule for setting interest rates; when $\alpha_\pi > 1$ the Taylor principle is said to be satisfied.

When, for example, $\alpha_x = 0$, the Taylor principle implies determinacy: there is a unique non-explosive solution to the system. Some authors have assumed that explosive solutions are not legitimate equilibria in this model, in which case determinacy provides for equilibrium uniqueness. Cochrane (2009, 2011) has argued prominently that in some cases explosive solutions are legitimate. His arguments on this point are most clearly made within a flexible-price environment. In this case prices adjust immediately to their market-clearing levels, which simplifies the system by eliminating the Phillips curve and making the output gap zero.

It is within the flexible-price environment that Cochrane and McCallum pursued their debate on whether the NK model uniquely identifies an equilibrium. McCallum (2009a,b) argued that the explosive paths are not learnable, regardless of whether they are legitimate equilibria. Cochrane (2009) on the other hand contended that McCallum's result hinged on the observability of exogenous shocks; indeed Cochrane claimed that when exogenous shocks are not observable the MSV solution is not stable

[17] 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.

under learning and the bubble (NF) solution *is* stable under learning. As we have discussed, we find compelling the argument that exogenous shocks are not observable. We now revisit this debate on learning and equilibrium selection using the tools we have developed. To make clear the connection between learning and temporary equilibrium outcomes we start with a micro-founded "agent-level" formulation of a flexible price economy, isomorphic to those used by Cochrane and McCallum.

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. There is a unit mass of households indexed as $\omega \in \Omega$. Each household aims to maximize discounted expected utility with discount factor $0 < \delta < 1$. Expectations are formed against subjective beliefs, and the utility flow is given by

$$u(c_t(\omega), m_{t-1}(\omega)) = \frac{1}{1-\sigma}\left(c_t(\omega)^{1-\sigma} - 1\right) + \log\left(\frac{m_{t-1}(\omega)}{\pi_t}\right),$$

with $\sigma > 0$. Here c is consumption, m is real balances and $\pi_t = p_t/p_{t-1}$ is the inflation factor. Note that $m_{t-1}(\omega) = M_{t-1}(\omega)/p_{t-1}$, where $M_{t-1}(\omega)$ is nominal money holdings at the end of time $t-1$, so that $m_{t-1}(\omega)/\pi_t = M_{t-1}(\omega)/p_t$ is the agent's real balances at the beginning of time t.

The budget constraint of the household is

$$c_t(\omega) + m_t(\omega) + b_t(\omega) = \frac{m_{t-1}(\omega)}{\pi_t} + \frac{R_{t-1}}{\pi_t} b_{t-1}(\omega) + \mathcal{E}_t(\omega) + \frac{\mathcal{T}_t}{p_t}, \tag{17}$$

where $\mathcal{E}_t(\omega)$ is household ω's time t real endowment of the perishable consumption good, R_{t-1} is the nominal interest rate factor, $b_t(\omega)$ 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 $\mathcal{E}_t(\omega) = 1$ for all $\omega \in \Omega$. Finally, we take as implicit that households cannot run Ponzi schemes.

The first-order conditions of household ω may be written

$$\begin{aligned} c_t(\omega)^{-\sigma} &= \delta R_t \hat{E}_t(\omega)\left(\frac{c_{t+1}(\omega)^{-\sigma}}{\pi_{t+1}}\right) &(18)\\ m_t(\omega) &= \delta\left(\frac{R_t}{R_t - 1}\right) c_t(\omega)^{\sigma}. &(19)\end{aligned}$$

Equation (18) is household ω's Euler equation capturing the trade-off between consumption and savings via risk-free claims, and equation (19) is money demand.

The government prints money and provides nominal injections (transfers). 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

$$M_t^s - M_{t-1}^s = \mathcal{T}_t. \tag{20}$$

Through transfers, the government can choose the nominal money supply M_t^s so that the nominal interest rate satisfies a Taylor rule:

$$R_t = 1 + \varphi_t f(\pi_t), \text{ where } f(\pi) = (R^* - 1)\left(\frac{\pi}{\pi^*}\right)^{\alpha_\pi\left(\frac{R^*}{R^*-1}\right)}. \tag{21}$$

Here φ_t captures a serially correlated policy shock, and 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 $\delta R^* = \pi^*$.

Market clearing provides the following restrictions:

$$\int_\Omega c_t(\omega)d\omega = \int_\Omega \mathcal{E}_t(\omega)d\omega = 1 \tag{22}$$

$$\int_\Omega m_t(\omega)d\omega = \frac{M_t^s}{p_t} \tag{23}$$

$$\int_\Omega b_t(\omega)d\omega = 0, \tag{24}$$

where the second equality in the first line follows from the assumption that $\mathcal{E}_t(\omega) = 1$ for all $\omega \in \Omega$.

To keep the analysis simple we assume homogeneity of expectations and of initial wealth, i.e. $b_{-1}(\omega) = 0$ for all ω. Because of this, in equilibrium all agents consume their endowment every period. To close the model we must specify precisely how consumption plans and expectations are formed. In contrast to the OLG example, here our agents are infinitely-lived, and there are alternative approaches to boundedly-rational decision-making. We adopt the Euler-equation learning approach for its simplicity and because it aligns with the one-step ahead framework used in Sections 2 and 3 and which was the focus of the debate between McCallum and Cochrane.[18]

We leverage this observation to assume that agents correctly forecast next period's consumption to equal the endowment. It follows therefore that agent ω's consumption $c_t(\omega) = c_t$ satisfies

$$c_t = \left(\delta R_t \hat{E}_t\left(\pi_{t+1}^{-1}\right)\right)^{-\frac{1}{\sigma}}, \tag{25}$$

and following the Euler-equation learning approach, we treat (25) as a behavioral primitive, giving the consumption demand of agent ω. The corresponding nominal money demand equation for agent ω is $M_t(\omega) = M_t$, where

$$M_t = p_t\left((R_t - 1)\hat{E}_t\left(\pi_{t+1}^{-1}\right)\right)^{-1}. \tag{26}$$

[18]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 (2015) to be asymptotically optimal. See also Evans and Honkapohja (2006).

Given φ_t, p_{t-1}, and M^s_{t-1}, temporary equilibrium values $(p_t, \pi_t, R_t, c_t, M_t, M^s_t, \mathcal{T}_t)$ are determined by the following equations:

$$\begin{aligned}
R_t &= 1 + \varphi_t f(\pi_t) \\
M_t &= p_t \left((R_t - 1)\, \hat{E}_t \left(\pi^{-1}_{t+1} \right) \right)^{-1} = M^s_t \\
c_t &= \left(\delta R_t \hat{E}_t \left(\pi^{-1}_{t+1} \right) \right)^{-\frac{1}{\sigma}} = 1 \\
M^s_t &= \mathcal{T}_t + M^s_{t-1} \\
\pi_t &= p_t / p_{t-1},
\end{aligned}$$

where expectations $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ are allowed to depend on current endogenous variables.[19] Bond market clearing is implied by Walras's Law, and homogeneity then implies $b_t(\omega) = 0$ for all ω.

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 $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$. For simplicity suppose that $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ is predetermined, i.e. independent of current p_t. The increase in the expectations term $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ 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 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^s_t via transfer payments.[20] To summarize, an exogenous reduction in expected inflation leads to lower inflation and interest rates in the temporary equilibrium. While our discussion presumed that $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ is predetermined, this is not necessary. If $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ 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 Sections 2 and 3, 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 (25) and the Taylor rule (21) yields

$$c_t = -\frac{1}{\sigma} \left(R_t - \hat{E}_t \pi_{t+1} \right) \tag{27}$$

$$R_t = \alpha_\pi \pi_t + \left(1 - \frac{\delta}{\pi^*} \right) \varphi_t, \tag{28}$$

[19] More generally $\hat{E}_t \left(\pi^{-1}_{t+1} \right)$ may depend on aditional lags of variables.

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

where all variables are now written as proportional deviations from means. Because in equilibrium $c_t = 0$, equations (27)–(28) simply reduce to

$$\pi_t = \frac{1}{\alpha_\pi} \hat{E}_t \pi_{t+1} + v_t \tag{29}$$

where v_t is the stationary AR(1) process given by $v_t = -\alpha_\pi{}^{-1}(1 - \delta\,(\pi^*)^{-1})\varphi_t$. Within a linearized framework this equation 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. In contrast, under adaptive learning a forecasting model is precisely specified, and given this forecasting model a unique equilibrium path is pinned down.[21]

The set-up is now in our standard form (1) and thus we can invoke our earlier results for the case in which the shocks are unobserved. By Theorem 1, if the Taylor principle is satisfied then the MSV solution is the unique robustly stable RE equilibrium under learning. This result may be interpreted as follows: regardless of the lag length of the agent's PLM, the temporary equilibrium path will converge to the MSV solution provided initial beliefs are not too far from the MSV coefficients; and further, the bubble solutions lack this critical stability property. 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 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 ... [W]e are finding the unique locally bounded equilibrium, not the unique equilibrium itself.

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. To see this, we return to the nonlinear model and suppose that $\varphi_t \equiv 1$ and that we are given p_{t-1} and M^s_{t-1}. Using (18) with $c_t(\omega) \equiv 1$, and the Taylor rule (21), it follows that there

[21] In the nonlinear environment above the precise forecasting model was not discussed; here, in our linear environment, agents base their forecasts $\hat{E}_t \pi_{t+1}$ on PLMs as discussed in Section 2.3.

is a continuum of perfect foresight paths $\{p_{t+n}\}_{n\geq 0}$, indexed by p_t, with $\pi_t = p_t/p_{t-1}$, satisfying

$$\pi_{t+n+1} = \delta\left(1 + f(\pi_{t+n})\right), \text{ for } n \geq 0.$$

Any such path identifies a legitimate equilibrium, with the values of the remaining variables given by

$$\begin{aligned}
R_{t+n} &= 1 + f(\pi_{t+n}), \\
p_{t+n} &= \pi_{t+n} p_{t+n-1}, \\
M_{t+n} &= \delta p_{t+n} R_{t+n} \left(R_{t+n} - 1\right)^{-1}, \\
\mathcal{T}_{t+n} &= M_{t+n} - M_{t+n-1}.
\end{aligned}$$

We note, in particular, that when $p_t > \pi^* p_{t-1}$ it follows that $\pi_t \to \infty$ along this equilibrium "bubble" path.

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. Under 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. When a rational expectations equilibrium is stable under adaptive learning, it can then be viewed as a emergent outcome of this process.[22]

In a related point, Cochrane suggests in various places 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. For example, 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?" As we have now clearly laid out, 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 identified. He suggests that this undermines the adaptive learning approach. However this misunderstands adaptive learning by private agents as requiring knowledge of the structural parameters. Under

[22]In related work Garcia Schmidt and Woodford (2015) examine the relevance of perfect foresight equilibria under an iterative process of belief revision. We regard our analysis and theirs as complementary approaches.

adaptive learning agents need to forecast in order to make decisions, but they do not generally need to know or estimate structural parameters in order to forecast. Instead they make forecasts the same way that time-series econometricians typically forecast: by estimating least-squares projections of the variables being forecasted on the relevant observables.

5 Conclusion

In the benchmark macroeconomic framework in which current aggregate endogenous variables are observable, rational expectations solutions are unaffected by whether or not the exogenous aggregate shocks driving those solutions are observable. This equivalence is established formally in Section 2 and holds both for determinate models, in which there is a unique non-explosive RE solution, and for indeterminate models with a multiplicity of such solutions. The observability issue is of importance because observability of key aggregate shocks to productivity and preferences, as well as to monetary policy, has forcefully been called into question.

An issue of central importance, since even the simplest forward-looking macroeconomic models have multiple RE solution paths, is the stability of the different REE under adaptive learning. As stressed by Cochrane (2011), when New Keynesian models are determinate, as when the interest rate rule satisfies the Taylor principle, RE solutions with explosive inflation paths (inflation bubbles) remain legitimate equilibria. While McCallum (2009) showed that the MSV solution, typically adopted by practitioners, is stable under learning, and the explosive inflation paths are not stable under learning, Cochrane (2009, 2011) argued that McCallum's results hinged on observability by agents of the exogenous shocks.

In Section 3 we addressed the critical issue of the stability of REE under adaptive learning when agents do not observe the exogenous aggregate shocks. We use standard adaptive learning tools, in which agents, like econometricians, are assumed to forecast key variables like inflation using a time-series forecasting model that is updated over time in response to observed data. Our key result is that MSV solutions are robustly stable under learning, while nonfundamental solutions, including inflation bubble solutions, are not robustly stable under learning.

In Section 4, which provides applications to the overlapping generations and New Keynesian models, we also addressed the important related issue raised by Cochrane of describing the causal mechanisms when agents forecast using adaptive learning. To do this we made explicit the temporary equilibrium setting in which market clearing occurs, building this up from agent-level decision making based on their forecasts of key aggregate variables. An equilibrium path with adaptive learning is then defined recursively: at each point in time market clearing delivers prices and quantities given agents' expectations; forecast-rule parameters are then updated and the revised rule is used to form expectations in the next period.

In concluding, we remark that using adaptive learning as a selection criterion to choose the MSV solution does preclude treating the adaptive learning dynamics themselves as of interest. Variations such as "constant gain" least-squares learning, in which past data is geometrically discounted, result in persistent learning dynamics around the MSV solution, and in some cases can include the possibility of temporary "escape" dynamics.[23] However, in these settings, the stable MSV solution remains the anchor for these extended learning dynamics.

[23] Constant gain learning dynamics have been emphasized, for example, in the New Keynesian model by Milani (2007) and in the RBC model by Eusepi and Preston (2010) and Mitra et al. (2013). Escape dynamics are a central focus in Cho, Sargent and Williams (2002), McGough (2006) and Branch and Evans (2011).

Appendix A: Proofs of Results

Proof of Proposition 1. While this result follows from, and corresponds to, the well-known Blanchard-Kahn conditions, it is worthwhile developing its simple argument within the context of Definition 1. Thus assume $|\beta| < 1$, and let beliefs $\mathcal{P}$ be a collection of beliefs identifying an equilibrium process y_t. Let

$$\eta_{t+1} = y_{t+1} - E^{T(\mathcal{P})}(y_{t+1}|y^t, v_t),$$

and note that η_t is a martingale difference sequence with respect to the distribution $T(\mathcal{P})$. Since

$$E^{T(\mathcal{P})}(y_{t+1}|y^t, v_t) = E^{\mathcal{P}}(y_{t+1}|y^t, v_t),$$

and since y_t must satisfy (3), we have

$$y_t = \beta^{-1}y_{t-1} + \eta_t - \beta^{-1}v_{t-1}. \tag{30}$$

Standard arguments show that if $\left|\beta^{-1}\right| > 1$ then y_t is explosive almost surely, unless $\eta_t = (1-\beta\rho)^{-1}\varepsilon_t$; in this latter case, the recursion (30) reduces to the MSV solution. This shows sufficiency. Necessity is simpler: if $|\beta| > 1$ then the NF solution satisfies the UBC. ■

Proof of Proposition 2. 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

$$\begin{aligned}
a &\to ((1-\rho)\gamma(\phi_1)\beta)\, a && (31)\\
\phi_1 &\to \rho + \beta\gamma(\phi_1)\phi_2 && (32)\\
\phi_n &\to \beta\gamma(\phi_1)(\phi_{n+1} - \rho\phi_n), \text{ for } n = 2, \ldots, N-1 && (33)\\
\phi_N &\to -(\beta\rho\gamma(\phi_1))\,\phi_N. && (34)
\end{aligned}$$

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 $N \geq 2$.

To show that all fixed points of T_N look like either ψ_N^{NF} or ψ_N^{MSV}, 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 $-\beta\rho\gamma(\phi_1^*) = 1$. If $\tilde{n} = N$ the we are done by (34). If $\tilde{n} < N$ then, since $\tilde{n} \geq 2$ and since $\phi_{\tilde{n}+1}^* = 0$, the claim follows by (33).

Having established $-\beta\rho\gamma(\phi_1^*) = 1$, we derive that $\phi_1^* = \beta^{-1} + \rho$. Then by (32)

$$\phi_2^* = \frac{\phi_1^* - \rho}{\beta\gamma(\phi_1^*)} = -\rho/\beta.$$

Finally, if $N \geq 3$ we may recursively compute, for $n \geq 3$, that

$$\phi_n^* = \left(\frac{1}{\beta\gamma(\phi_1^*)} + \rho \right) \phi_{n-1}^* = 0.$$

It follows that $\psi^* = \psi_N^{NF}$, completing the proof. ■

Proof of Theorem 1. We need to assess E-stability of the MSV and NF solutions for each N using (11) 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 (31) and T_N^ϕ denote the map given by (32)-(34). 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

$$\frac{\partial}{\partial a} T_N^a = \frac{(1-\rho)\beta}{1-\beta\rho}.$$

Also, $\partial T_1^\phi / \partial \phi_1 = 0$ and, for $N \geq 2$,

$$\frac{\partial}{\partial \phi'} T_N^\phi = \begin{pmatrix} 0 & \frac{\beta}{1-\beta\rho} & 0 & 0 & 0 & \cdots & 0 \\ 0 & -\frac{\beta\rho}{1-\beta\rho} & \frac{\beta}{1-\beta\rho} & 0 & 0 & \cdots & 0 \\ 0 & 0 & -\frac{\beta\rho}{1-\beta\rho} & \frac{\beta}{1-\beta\rho} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\beta\rho}{1-\beta\rho} & \frac{\beta}{1-\beta\rho} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\frac{\beta\rho}{1-\beta\rho} \end{pmatrix},$$

which yields eigenvalues of 0 and $-\frac{\beta\rho}{1-\beta\rho}$. For E-stability we need $-\frac{\beta\rho}{1-\beta\rho} < 1$ and $\frac{(1-\rho)\beta}{1-\beta\rho} < 1$. 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.

We now turn our attention to ψ_N^{NF}. We have

$$\frac{\partial}{\partial a} T_N^a = -\frac{1-\rho}{\rho},$$

which satisfies $\frac{\partial}{\partial a} T_N^a < 1$ since $0 < \rho < 1$. We next compute

$$\frac{\partial}{\partial \phi'} T_N^\phi - I_N = \begin{pmatrix} -\frac{1}{\beta\rho} - 1 & -\frac{1}{\rho} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{\beta} & 0 & -\frac{1}{\rho} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -\frac{1}{\rho} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\rho} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{35}$$

where I_N is the $N \times N$ identity matrix. A necessary condition for E-stability is that the eigenvalues of (35) be less than or equal to zero. If $N = 2$, then the eigenvalues are -1 and $-\frac{1}{\beta\rho}$. If $N > 2$ then the eigenvalues are just as with $N = 2$ plus a zero eigenvalue with multiplicity $N - 2$. 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 (11) 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.

Diagonalize the derivative of F as $DF(0) = S\Lambda S^{-1}$:

$$\begin{pmatrix} -\frac{1}{\beta\rho} - 1 & -\frac{1}{\rho} & 0 \\ \frac{1}{\beta} & 0 & -\frac{1}{\rho} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -\beta & -\frac{1}{\rho} & \frac{\beta}{\rho} \\ 1 & 1 & -\frac{\beta\rho+1}{\rho} \\ 0 & 0 & 1 \end{pmatrix}$$

$$\times \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{\beta\rho} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\rho}{1-\beta\rho} & \frac{1}{1-\beta\rho} & \frac{1}{\rho-\beta\rho^2} \\ \frac{\rho}{\beta\rho-1} & \frac{1}{\beta\rho-1} + 1 & \beta\left(\frac{1}{\beta\rho-1} + 1\right) \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, 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.

Now change coordinates: $\zeta = S^{-1}\varphi$. It follows that

$$\begin{aligned} d\zeta/d\tau &= S^{-1}d\varphi/d\tau = S^{-1}F(\varphi) && (36) \\ &= S^{-1}DF(0)\varphi + S^{-1}\hat{H}(\varphi) \\ &= \Lambda\zeta + S^{-1}\hat{H}(S\zeta), \text{or} \\ d\zeta/d\tau &= \Lambda\zeta + H(\zeta), && (37) \end{aligned}$$

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 (37).

The system (37) has two stable eigenvalues and one zero eigenvalue. By the center-manifold theorem there is a neighborhood of the origin, $U \subset \mathbb{R}^3$, and one-dimensional, smooth "center" manifold $\mathcal{W}_c \subset U$, so that the following hold:

1. The manifold $\mathcal{W}_c$ is an invariant subspace of the dynamic system (37).
2. There exists an open set $V \subset \mathbb{R}$ containing the origin, and smooth real-valued functions $h_i : V \to \mathbb{R}$, for $i = 1, 2$ so that
$$\mathcal{W}_c = \{(h_1(\zeta_3), h_2(\zeta_3), \zeta_3) : \zeta_3 \in V\}.$$
3. The functions h_i are zero to first order: $h_i(0) = h_i'(0) = 0$ for $i = 1, 2$.
4. The local dynamics of the system (37) are asymptotically equivalent to dynamics of system (37) restricted to $\mathcal{W}_c$.

Items 2 and 4 imply that the origin in $\mathbb{R}^3$ is a Lyapunov-stable steady state of (37) if and only if the origin in $\mathbb{R}$ is an Lyapunov-stable steady state of

$$d\zeta_3/d\tau = H^3\left(h_1(\zeta_3), h_2(\zeta_3), \zeta_3\right), \tag{38}$$

where a superscript indicates the coordinate in the range, i.e. $H = (H^1, H^2, H^3)'$.

We are required now to determine the explicit form of H^3. Since the last row of S^{-1} is $(0, 0, 1)$, by equation (36) 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 (37) 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

$$\begin{aligned} F^3(\varphi) &= T_3^{\phi,3}(\varphi + \phi^*) - (\varphi + \phi^*) \\ &= -\frac{\beta\rho}{1 - \beta(\varphi_1 + \phi_1^*)}\varphi_3 - \varphi_3 \\ &= \frac{\rho\varphi_3}{\varphi_1 + \rho} - \varphi_3 \\ &= -\frac{\varphi_1\varphi_3}{\varphi_1 + \rho}, \end{aligned}$$

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

$$\begin{aligned} H^3(\zeta) = F^3(S\zeta) &= -\frac{\zeta_3(-\beta\zeta_1 - \rho^{-1}\zeta_2 + \beta\rho^{-1}\zeta_3)}{\rho - \beta\zeta_1 - \rho^{-1}\zeta_2 + \beta\rho^{-1}\zeta_3}, \text{ or} \\ H^3(\zeta) &= \frac{\zeta_3(\zeta_2 - \beta(\zeta_3 - \rho\zeta_1))}{\beta(\zeta_3 - \rho\zeta_1) - \zeta_2 + \rho^2}. \end{aligned} \tag{39}$$

We conclude that

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

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 may compute

$$\begin{aligned} G' &= H_1^3 h_1' + H_2^3 h_2' + H_3^3 \\ G'' &= (H_{11}^3 h_1' + H_{12}^3 h_2' + H_{13}^3) h_1' + H_1^3 h_1'' \\ &+ (H_{21}^3 h_1' + H_{22}^3 h_2' + H_{23}^3) h_1' + H_2^3 h_2'' \\ &+ H_{31}^3 h_1' + H_{32}^3 h_2' + H_{33}^3, \end{aligned}$$

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 (39) it is immediate the $H_i^3 = 0$ for $i = 1, 2, 3$. Also, by item 3 above, we have that $h_i = h_i' = 0$ for $i = 1, 2$. It follows, then, that $G' = 0$ and $G'' = H_{33}^3$. Again using (39), we compute directly that $G'' = -\frac{2\beta}{\rho^2}$.

The above computations allow us to use Taylor's theorem to write

$$d\zeta_3/d\tau = -\frac{\beta}{\rho^2}\zeta_3^2 + \mathcal{O}(|\zeta_3|^3).$$

It follows that the origin is not a Lyapunov stable fixed point of (38): there exists $\varepsilon, \delta > 0$ so that if the system (38) 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 (11).

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

$$\frac{d}{d\tau}\phi(N) = T_N^{\phi}(\phi(N)) - \phi(N),$$

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 the T-map, notice that

$$d\phi_N(N)/d\tau = \left(-\beta\rho\gamma(\phi_1(N)) - 1\right)\phi_N(N).$$

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}:

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

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, (32)-(33), 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: Nonfundamental Equilibria

Recall the dynamic system (1)-(2), rewritten here for convenience:

$$\begin{aligned} y_t &= \beta E_t y_{t+1} + v_t \\ v_t &= \rho v_{t-1} + \varepsilon_t. \end{aligned} \tag{40}$$

We are interested in solutions to (40) that are consistent with PLM (4), also rewritten here for convenience:

$$y_t = a + \sum_{i=1}^{N} \phi_i y_{t-i} + \sum_{j=0}^{M} \lambda_j v_{t-j}. \tag{41}$$

We have the following result:

Proposition 3 *The process y_t is a solution to (40) consistent with PLM (41) if and only if*

$$y_t = \beta^{-1} y_{t-1} - \left(\beta^{-1} + \theta\rho\right) v_{t-1} + \theta v_t \tag{42}$$

for some $\theta \in \mathbb{R}$.

Proof. If y_t satisfies

$$y_t = \beta^{-1} y_{t-1} - \beta^{-1} v_{t-1} + \eta_t \tag{43}$$

for some martingale difference sequence η_t then y_t is solution to (40). Now let $\theta \in \mathbb{R}$, and set

$$\eta_t = \theta\varepsilon_t = \theta\left(v_t - \rho v_{t-1}\right). \tag{44}$$

It follows that

$$y_t = \beta^{-1} y_{t-1} - \left(\beta^{-1} + \theta\rho\right) v_{t-1} + \theta v_t.$$

Thus any process y_t satisfying (42) is a solution to (40) consistent with PLM (41). Now suppose y_t is solution to (40) that is consistent with PLM (41). Then

$$\lambda_0 \varepsilon_t = y_t - E_{t-1} y_t = \eta_t,$$

where the second equality follows from (43). Setting $\theta = \lambda_0$ and using (44) completes the argument.■

By Proposition 3, every solution to (40) consistent with PLM (41) takes the form (42), and thus the MSV and NF solutions must correspond to particular values of θ. When $\theta = -(\beta\rho)^{-1}$, the NF solution is recovered, and when $\theta = (1-\beta\rho)^{-1}$, canceling the common factor $1 - \beta^{-1}L$ provides the MSV solution.

Even though there are many non-fundamental solutions to (40) consistent with PLM (41), our NF solution is unique in that it has an autoregressive representation of finite order. Specifically, we have the following result:

Proposition 4 *If y_t is solution to (40) consistent with PLM (41) and if*

$$y_t = \sum_{n=1}^{p} \varphi_n y_{t-n} + \mu_t, \tag{45}$$

for some innovations μ_t, then y_t is either the MSV or the NF solution.

Proof. If y_t satisfies (40), (41) and (45) then $\eta_t = \lambda_0 \varepsilon_t = y_t - E_{t-1} y_t = \mu_t$. Writing $v_{t-1} = \rho^{-1}(v_t - \varepsilon_t)$ and using (43) and that $v_t = (1-\rho L)^{-1} \varepsilon_t$, we have

$$\begin{aligned}
y_t &= \beta^{-1} y_{t-1} - (\beta\rho)^{-1} (v_t - \varepsilon_t) + \eta_t \\
&= \beta^{-1} y_{t-1} + (\beta\rho)^{-1} \varepsilon_t - (\beta\rho)^{-1} (1-\rho L)^{-1} \varepsilon_t + \eta_t \\
&= \left(\beta^{-1} + \rho\right) y_{t-1} - \beta^{-1}\rho y_{t-2} - (\beta\rho)^{-1} \varepsilon_t + (1-\rho L) \left(\eta_t + (\beta\rho)^{-1} \varepsilon_t\right), \text{ or} \\
y_t &= \left(\beta^{-1} + \rho\right) y_{t-1} - \beta^{-1}\rho y_{t-2} + \mu_t - \left(\rho + \beta^{-1}\lambda_0^{-1}\right) \mu_{t-1},
\end{aligned} \tag{46}$$

where the third equality follows from acting on each side by $1 - \rho L$, and the fourth by replacing η_t and ε_t with μ_t and $\lambda_0^{-1}\mu_t$, respectively. For (45) and (46) to be consistent, it must be that $p \leq 2$. If $p = 2$ then $\rho + \beta^{-1}\lambda_0^{-1} = 0$. Solving, we obtain $\lambda_0 = -(\beta\rho)^{-1}$, which provides the NF solution. If $p = 1$ then there must be a common factor which cancels the second lag of y. Equation (46) may be written

$$(1 - \beta^{-1}L)(1-\rho L) y_t = \left(1 - \left(\rho + \beta^{-1}\lambda_0^{-1}\right) L\right) \mu_t.$$

Since $\rho + \beta^{-1}\lambda_0^{-1} \neq \rho$, we consider only the case in which $\rho + \beta^{-1}\lambda_0^{-1} = \beta^{-1}$. Solving, we obtain $\lambda_0 = (1-\beta\rho)^{-1}$, which is the MSV solution.■

In this paper we focused on solutions to (40) that are consistent with the PLM (41). By Proposition 3 any such solution takes the form (42). In Section 2.2 we identified two particular solutions of interest: the MSV solution and the NF solution. By Proposition 4 we established that these are precisely the solutions to (40) that are both consistent with the PLM (41) and allow for finite order autoregressive (AR(p)) representations. The existence of AR(p) representations considerably simplifies the asymptotic analysis of learning dynamics when the exogenous shocks are taken as not observed because the recursive least-squares estimator may be applied. For general $\theta \in \mathbb{R}$, solutions of the form (42) have ARMA(2,1) representations. The corresponding recursive estimators are more involved and additional technical apparatus is required for the asymptotic analysis. See Evans and Honkapohja (1994) for details.

References

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

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

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

[4] Cho, I.-K., Williams, N. and Sargent, T. (2002), Escaping Nash inflation. Review of Economic Studies 69, 1-40.

[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] Eusepi, S. and Preston, B. (2010), Expectations, learning and business cycle fluctuations. American Economic Review 101, 2844-2872.

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

[9] Evans and Honkapohja, S. (1994), Learning, convergence, and stability with multiple rational expectations equilibria. European Economic Review 38, 1071-1098.

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

[11] Evans, G. and Honkapohja, S. (2003), Expectations and the stability problem for optimal monetary policy. Review of Economic Studies 70, 807-824.

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

[13] Evans, G. and Honkapohja, S. (2009), Learning and macroeconomics, Annual Review of Economics 1, 421-449.

[14] Evans, G. and McGough, B. (2015), Learning to optimize, Mimeo.

[15] Garcia Schmidt, M. and Woodford, M. (2015), Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis, presentation at the NBER Summer Institute, July 9, 2015.

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

[17] King, Robert G. and Rebelo, S. (1999), Resuscitating real business cycles, in: J. B. Taylor and M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 14, 927-1007 Elsevier.

[18] 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.

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

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

[21] 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.

[22] 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.

[23] 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.

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

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

[26] McGough, B. (2006), Shocking escapes. Economic Journal 116, 507-528.

[27] Milani, F. (2007), Expectations, learning and economic persistence. Journal of Monetary Economics 54, 2065-2082.

[28] Mitra, K., Evans, G. and Honkapohja, S. (2013), Policy change and learning in the RBC model. Journal of Economic Dynamics and Control 37, 1947-1971.

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

[30] Sims, C. (2003), Implications of rational inattention. Journal of Monetary Economics 50, 665-690.

[31] Taylor, J. and Woodford, M. (ed.), (1999), Handbook of Macroeconomics, edition 1, volume 1, chapter 14, 927-1007 Elsevier.

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