

LEARNING AND INFORMATIONAL STABILITY
OF DYNAMIC REE WITH INCOMPLETE INFORMATION*

Giacomo Rondina[†] Todd B. Walker[‡]

March 2015

ABSTRACT

In the context of dynamic models of incomplete information, we show that slight perturbations to the agents' information sets can lead to vastly different Rational Expectations Equilibria (REE). The difference is due to a hidden instability (i.e., an exact cancellation of an explosive autoregressive root) that is a property of the full-information equilibrium but not present in the partial-information equilibrium. Due to the multitude of potential equilibria, we use least-squares learnability as a refinement mechanism. We find that models with complete information revelation are not least-squares learnable, while information structures that do not fully reveal the underlying shocks are learnable. We show that learnability relates to the informational stability properties of an equilibrium, whereby an equilibrium is said to be informationally unstable if it vanishes when information is slightly perturbed. We present application to a model with productivity shocks and nominal rigidities. In both cases we show that equilibria with complete information are informationally unstable, and thus not learnable; while equilibria that preserve incomplete information are informationally stable and learnable.

Keywords: Rational Expectations, Incomplete Information, Least-Squares Learning, E-Stability

*We would like to thank an associate editor and three anonymous referees for helpful comments. We would also like to thank Marios Angeletos, Tim Cogley, Thorsten Drautzburg, Thomas Mertens, Gary Ramey, Venky Venkateswaran, and Mirko Wiederholt for useful conversations. We acknowledge financial support from the National Science Foundation under grant number SES-0962221. Part of this research was conducted while visiting the Laboratory for Aggregate Economics and Finance at UC-Santa Barbara.

[†]University of Colorado, giacomo.rondina@colorado.edu

[‡]Indiana University, walkertb@indiana.edu

1 INTRODUCTION

Within the context of a stylized but quite general model, we show that slight perturbations to agents' information sets can lead to vastly different Rational Expectations equilibria. Due to the multitude of potential equilibria, we use least-squares learnability as an equilibrium refinement mechanism. Our main theorem shows that models in which the endogenous variables reveal perfectly the unobservable state of the world are not robust to small perturbations in the information set of the agents, and are therefore not learnable. However, models in which incomplete information persists in equilibrium can be learned using a least-squares algorithm.

Borrowing from the engineering literature, we show that the reason complete information equilibria cannot be learned is due to a “hidden instability,” which is an exact cancellation of an explosive autoregressive root. The cancellation works only along the equilibrium path. When information is perturbed, the hidden instability emerges and the process becomes explosive. For this reason, least-squares algorithms will not converge. This hidden instability does not emerge in partial-information equilibria. Perturbations to agents' information sets in this case do not lead to explosive dynamics and these equilibria are least-squares learnable. We show that this is due to how expectations are formed when information is incomplete.

This type of instability and connection to the learning literature is, to our knowledge, novel. There has been a recent resurgence in dynamic models with incomplete information, but the question of robustness of these equilibria has yet to be fully studied.¹ We follow a rich history of papers that use least-squares learning as an equilibrium refinement criteria [see Evans and Honkapohja (2001)]. However, we are unaware of any papers that focus exclusively on learning in models of incomplete information.

It has been argued that learnable equilibria are more likely to occur in reality and therefore learnability is an important criterion for rational expectations models [Evans and Honkapohja (2001)]. In light of our results, we argue that more emphasis should be placed on the learnability of incomplete-information, rational expectations equilibria. While our results pertain to a particular class of models, this paper lays out the framework for which that can be accomplished.

Finally, we present an application of our results. We analyze a stylized macroeconomic model with nominal rigidities and permanent productivity shocks similar to Lorenzoni (2009). We show that the fully revealing rational expectations equilibrium of that model is not learnable when the productivity process takes the form of an S-shaped diffusion process. We then solve for an incomplete information rational expectations equilibrium of the model and show that it is learnable. We analyze the impulse response of output and inflation to productivity innovations and we show that in the incomplete information equilibrium a positive productivity innovation leads to alternating

¹Recent examples include Woodford (2003a), Pearlman and Sargent (2005), Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), Hellwig (2006), Adam (2007), Gregoir and Weill (2007), Angeletos and Pavan (2007), Kasa, Walker, and Whiteman (2011), Lorenzoni (2009), Rondina (2009), Angeletos and La'O (2009), Angeletos and La'O (2011), Hellwig and Venkateswaran (2009), Graham and Wright (2010), Nimark (2011), Hassan and Mertens (2011). Kasa (2000) contains an E-stability result for the model of Townsend (1983).

oscillations of output, employment and inflation around their respective trends. More specifically, both output and inflation are above trend and below trend at the same time, so that, conditional on the productivity innovation output and inflation display a positive correlation over the cycle. In other words, for the learnable equilibrium, a positive supply shock (in productivity) generates a dynamic response that looks like alternating positive and negative demand shocks.

2 MODEL, INFORMATION, AND EQUILIBRIUM

We perform the analysis in the simplest possible setting to make clear the issues at hand.² In particular, we focus on a univariate model with a representative agent facing incomplete information about the state of the economy. The equilibrium equation is specified as

$$y_t = \kappa \mathbb{E}[y_{t+1} | \Omega_t] + \varphi a_t \quad (2.1)$$

with $|\kappa| \leq 1$ and $\varphi \in \mathbb{R}$. Let Ω_t denote the information set of the representative agent at time t , to be specified shortly, and a_t is an exogenous stochastic process given by

$$a_t - \rho a_{t-1} = A(L)u_t \quad (2.2)$$

where $|\rho| \leq 1$, $A(L)$ is a square summable lag polynomial in non-negative powers of L and $u_t \sim N(0, \sigma_u^2)$. The immediate consequence of working in a linear-quadratic Gaussian setting is that the expectational operator corresponds to the linear projection operator, a property that will apply to all results in the paper. In what follows we restrict our attention to non-explosive solutions to (2.1).

As is standard in the rational expectations literature, we assume that agents' information comes from two sources—exogenous and endogenous. Exogenous information, denoted \mathbb{U}_t , is that which is not generated by equilibrium forces but is revealed to the agents from outside of the context of the model (e.g., an exogenous signal about u_t). Endogenous information, denoted $\mathbb{V}_t(y)$, comes from the knowledge that y_t is generated by (2.1). In our context this results in the entire history of a_t always being part of the equilibrium information of the representative agent. This is immediately evident by considering that, along any equilibrium y_t , the representative agent observes the prediction $\mathbb{E}[y_{t+1} | \mathbb{V}_t(y)]$ and so she must be able to compute

$$a_t = \frac{1}{\varphi} (y_t - \kappa \mathbb{E}[y_{t+1} | \mathbb{V}_t(y)]) \quad (2.3)$$

Throughout our analysis we thus take for given that $\mathbb{V}_t(a) \subseteq \mathbb{V}_t(y)$.

An important benchmark for the following analysis is the rational expectations equilibrium of (2.1) when the representative agent is (exogenously) endowed with the knowledge of the entire history of shocks up to time t . We refer to this equilibrium as the Full Information equilibrium.

²Appendix B provides conditions under which the results derived below extend to more complex models.

Let $\mathbb{V}_t(x)$ denote the smallest closed linear subspace spanned by the infinite history of the random variable x_t up to time t , namely $x^t \equiv \{x_t, x_{t-1}, x_{t-2}, \dots\}$. In a full information equilibrium it is assumed that $\Omega_t = \mathbb{V}_t(u)$. The rational expectations equilibrium for the full information case is well known and given by

$$y_t - \rho y_{t-1} = \frac{\varphi}{L - \kappa} \left(LA(L) - \kappa A(\kappa) \frac{(1 - \rho L)}{(1 - \rho \kappa)} \right) u_t \quad (2.4)$$

If we altered the information structure so that the agents could not condition directly on the structural shocks but only saw the current and past realizations of endogenous variables, $\Omega_t = \mathbb{V}_t(y)$, under what conditions would the equilibrium be fully revealing? That is, when would the span of $\mathbb{V}_t(y)$ be equivalent, in mean-square norm, to the span of $\mathbb{V}_t(u)$?

Since $\mathbb{V}_t(a) \subseteq \mathbb{V}_t(y)$, the answer to this question depends on the properties of $A(L)$, the exogenous stochastic process. The key feature will be the invertibility of the process, that is to say whether the observation of the history of a_t is able to perfectly reveal the history of the innovations u_t . This property is related to the moving average roots of $A(L)$. If the roots are all outside the unit circle, then $A(L)$ is said to be invertible in current and past a_t , and the history of a_t reveals perfectly the history of u_t . On the other hand, if at least one root is inside the unit circle, then $A(L)$ is non-invertible in current and past a_t , and the history of a_t is able to reveal only an imperfect measure of u_t .

When $A(L)$ is non-invertible, a_t can always be written as

$$a_t - \rho a_{t-1} = \tilde{A}(L) \prod_{i=1}^m (1 + \theta_i L) u_t, \quad |\theta_i| > 1 \text{ for } i = 1, \dots, m, \quad (2.5)$$

where $m > 0$ is the number of roots inside the unit circle and $\tilde{A}(L)$ is the invertible portion of $A(L)$. Restricting our attention to the case of $m = 1$, with $\theta \equiv \theta_1$ and $\rho = 0$, equation (2.5) can be expanded as³

$$\sum_{j=0}^{t-1} (-\theta)^j a_{t-j} = u_t + (-\theta)^t u_0. \quad (2.6)$$

The process a_t is invertible when $|\theta| < 1$, which implies that as t grows large, the summation on the left hand side of (2.6) remains well defined and is exactly equal to u_t since the term $(-\theta)^t u_0$ tends to zero for any finite u_0 . When $|\theta| > 1$, a_t is not invertible and so as t grows larger and larger the summation on the left hand side does not have a well defined limit. If the structural innovation, u_0 is not observed, which is the assumption under which (2.6) is valid, the exact knowledge of the structural innovation will never be revealed. Thus, insofar as agents cannot infer the structural innovations directly, non-invertible representations can be used to formulate incomplete information

³Allowing for an arbitrary ρ is immediate, just substitute any instance of a_t by the quasi-difference $a_t - \rho a_{t-1}$.

models [Futia (1981), Townsend (1983), Kasa (2000)].

When the process is non-invertible and u_t is not in the exogenous information set, the agents must solve a filtering problem. The information contained in the history of a_t in presence of non-invertibility can be summarized by its Wold fundamental representation, which is

$$a_t - \rho a_{t-1} = \tilde{A}(L) \prod_{i=1}^m (\theta_i + L) \tilde{u}_t, \quad (2.7)$$

$$\tilde{u}_t \equiv \prod_{i=1}^m \left(\frac{1 + \theta_i L}{\theta_i + L} \right) u_t \quad (2.8)$$

The innovation process \tilde{u}_t is the measure that minimizes the mean squared forecast error in predicting a_{t+1} linearly from its past. It is related to u_t via the Blaschke factor, $(1 + \theta_i L)/(\theta_i + L)$.⁴ The information contained in \tilde{u}_t is less than that of u_t . This is seen by noting that contemporaneous \tilde{u}_t is a linear combination of *past* u_t 's. Expanding (2.7) with $\rho = 0$ and $m = 1$ yields

$$\sum_{j=0}^{t-1} (-\theta)^{-j} a_{t-j} = \tilde{u}_t + (-\theta)^{-t} u_0, \quad (2.9)$$

and as t becomes arbitrarily large, the sum on the left hand side remains well defined for $|\theta| > 1$, implying that the impact of the initial state u_0 dissipates over time. However, the process reveals \tilde{u}_t , which is a strict subset of the structural innovations u_t .

The following proposition characterizes the rational expectations equilibrium for alternative information structures and specifications of the exogenous process.

Proposition 1. Rational Expectations Equilibria

Case [i.] If $m = 0$ ($A(L)$ is invertible), the Full Information equilibrium given by (2.4) is the Rational Expectations Equilibrium, independent of the exogenous information structure \mathbb{U}_t .

Case [ii.] If $m > 0$ ($A(L)$ is not invertible), the Full Information equilibrium (2.4) is the Rational Expectations Equilibrium from time $t > 0$ onward if $\{u_{t-j}, u_{t-j-1}, \dots, u_{t-j-m+1}\} \in \mathbb{U}_t$ with $0 < j < \infty$. If $j = \infty$ the Full Information equilibrium is always the Rational Expectations Equilibrium.

Case [iii.] If $m > 0$ ($A(L)$ is not invertible) and the exogenous information is specified as $\mathbb{U}_t = \{\emptyset\}$ for $t \in \mathbb{Z}$, then the stochastic process

$$y_t - \rho y_{t-1} = \frac{\varphi}{L - \kappa} \left\{ LA(L) - \kappa A(\kappa) \frac{(1 - \rho L) \prod_{i=1}^m \mathcal{B}_{\theta_i}(L)}{(1 - \rho \kappa) \prod_{i=1}^m \mathcal{B}_{\theta_i}(\kappa)} \right\} u_t \quad (2.10)$$

where

$$\mathcal{B}_{\theta_i}(L) \equiv \frac{1 + \theta_i L}{\theta_i + L},$$

⁴See Hansen and Sargent (1991) and Lippi and Reichlin (1994) for more on Blaschke factors.

is the unique stationary rational expectations equilibrium of (2.1).

Proof. See Appendix A. □

Case [i.] is trivial. Case [ii.] states that if $A(L)$ is not invertible, then the structural innovations must be included in the agents' information set directly via the exogenous information \mathbb{U}_t in order for the full information equilibrium to be the rational expectations equilibrium. Case [ii.] also distinguishes between that of an arbitrary long, yet finite, history, from that of a non-finite history. This is useful for thinking about cases in which the state was exactly observed at some distant point in the past.⁵ Case [iii.] gives the equilibrium in which information remains incomplete. The agents' filtering problem delivers the structural shocks of (2.7), and by virtue of cross-equation restrictions, these filtered innovations enter the endogenous variable according to (2.10).

Thus we have established that there are two rational expectations equilibria – one that is fully revealing (2.4), and one in which incomplete information is preserved (2.10). Moreover, we demonstrate in Section 4 that the dynamic properties of these equilibria are *markedly* different. Given that there are two rational expectations equilibria that (a priori) are equally likely, we follow Evans and Honkapohja (2001), among others, and propose least-squares learnability as a refinement mechanism.

3 LEARNING AND INCOMPLETE INFORMATION

3.1 MAIN RESULT Our learning analysis follows the methods and the notation of Chapter 6 in Evans and Honkapohja (2001). The first step in the learning analysis is to recast the equilibrium relationships in terms of a bi-variate process for the equilibrium outcome y_t and the implied forecast errors, which we will denote by e_t .⁶ Formally, for any rational expectations information equilibrium define the forecast error process as $e_t \equiv y_t - \mathbb{E}[y_t | \mathbb{V}_{t-1}(y) \vee \mathbb{U}_{t-1}]$, which is the span of past endogenous variables and exogenous information.

Without loss of generality, we will focus on the case of $a_t = u_t + \theta u_{t-1}$ with $\theta > 1$.⁷ Under this specification for a_t , both rational expectations equilibria can be then represented as

$$y_t = \kappa \eta e_t + \varphi a_t \tag{3.1}$$

$$e_t = y_t - \eta e_{t-1}, \tag{3.2}$$

⁵Interestingly, this relates to the approach of “truncating” the informational incompleteness by assuming the current state is revealed to the agents at some point in the future, an assumption that simplifies the characterization of the rational expectations solution and that is widely used in the incomplete information literature since Townsend (1983) first proposed it. Recent examples of papers that employ the truncation approach include Hellwig (2002) and Lorenzoni (2009). For a formal treatment of truncation see Nimark (2011). In our context the horizon of truncation is inconsequential for the equilibrium dynamics: as long as the state is eventually revealed, the full information of Case [ii] applies.

⁶We are thankful to Pierre-Olivier Weill for suggesting to us to work within the bi-variate representation that includes the process of forecast errors.

⁷The results carry through with $\rho \in (0, 1)$ and $\tilde{A}(L)$ invertible. The case of negative θ follows the same steps but it would make notation more burdensome.

where the coefficient η takes different forms depending on the equilibrium under consideration. More precisely, denote by $\hat{\eta}$ the coefficient for Full Information equilibrium (2.4), and by η^* the coefficient for the Incomplete Information equilibrium (2.10), where

$$\hat{\eta} = \frac{\theta}{1 + \kappa\theta} \quad \text{and} \quad \eta^* = \frac{1}{\kappa + \theta}. \quad (3.3)$$

We focus on least-squares learning of the parameter η and specify the learning algorithm following Evans and Honkapohja (2001)

$$y_t = \kappa\eta_{t-1}e_t + \varphi a_t \quad (3.4)$$

$$e_t = y_t - \eta_{t-1}e_{t-1} \quad (3.5)$$

$$\eta_t = \frac{\frac{1}{t} \sum_{s=1}^t y_s e_{s-1}}{\frac{1}{t} \sum_{s=1}^t e_{s-1}^2}, \quad (3.6)$$

with η_0 and e_0 given. This learning system can be written in recursive form as follows

$$e_t = -\frac{\eta_{t-1}}{1 - \kappa\eta_{t-1}}e_{t-1} + \frac{\varphi}{1 - \kappa\eta_{t-1}}(u_t + \theta u_{t-1}) \quad (3.7)$$

$$\eta_t = \eta_{t-1} + \frac{1}{t} S_{t-1}^{-1} e_t e_{t-1} \quad (3.8)$$

$$S_t = S_{t-1} + \frac{1}{t}(e_t^2 - S_{t-1}) + \frac{1}{t^2} \frac{-t}{t+1}(e_t^2 - S_{t-1}), \quad (3.9)$$

where S_t represents the time- t estimate of the variance-covariance matrix of the process for the forecast errors. To apply the formal results of Chapter 6 in Evans and Honkapohja (2001) it is useful to represent the system in compact notation, so we define

$$X_t = \begin{pmatrix} e_t \\ e_{t-1} \\ u_t \end{pmatrix} \quad \text{and} \quad \lambda_t = \begin{pmatrix} \eta_t \\ S_t \end{pmatrix}. \quad (3.10)$$

In the two equilibria that we are considering the vector λ takes the form

$$\hat{\lambda} = \begin{pmatrix} \hat{\eta} \\ \varphi^2(1 + \kappa\theta)^2 \end{pmatrix}, \quad \text{and} \quad \lambda^* = \begin{pmatrix} \eta^* \\ \varphi^2(\kappa + \theta)^2 \end{pmatrix}. \quad (3.11)$$

The learning algorithm can then be written as

$$X_t = G(\lambda_{t-1})X_{t-1} + F(\lambda_{t-1})u_t \quad (3.12)$$

$$\lambda_t = \lambda_{t-1} + \gamma_t H(\lambda_{t-1}, X_t) + \gamma_t^2 \rho_t(\lambda_{t-1}, X_t), \quad (3.13)$$

where

$$\begin{aligned}
G(\lambda_{t-1}) &\equiv \begin{pmatrix} -\frac{\eta_{t-1}}{1-\kappa\eta_{t-1}} & 0 & \frac{\theta\varphi}{1-\kappa\eta_{t-1}} \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & F(\lambda_{t-1}) &\equiv \begin{pmatrix} \frac{\varphi}{1-\kappa\eta_{t-1}} \\ 0 \\ 1 \end{pmatrix}, \\
\gamma_t &= \frac{1}{t}, & H(\lambda_{t-1}, X_t) &= \begin{pmatrix} S_{t-1}^{-1}e_t e_{t-1} \\ e_t^2 - S_{t-1} \end{pmatrix}, & \rho_t(\lambda_{t-1}, X_t) &= \begin{pmatrix} 0 \\ \frac{-t}{t+1}(e_t^2 - S_{t-1}) \end{pmatrix}.
\end{aligned}$$

An equilibrium λ can be learned if the above dynamic system converges to λ as $t \rightarrow \infty$ for λ_0 in a neighborhood of λ . For such convergence to happen several properties of the components of the dynamic system have to be satisfied. A necessary condition for convergence is that the process for X_t is stationary in a neighborhood of λ , which corresponds to the Eigenvalues of $G(\lambda)$ having modulus smaller than one.

The Eigenvalues of the matrix $G(\hat{\lambda})$ for the Full Information case are given by the vector $(-\theta, 0, 0)$. Since we have assumed that $\theta > 1$ and $G(\lambda)$ is continuous, it follows that the process X_t is non-stationary in a neighborhood of $\hat{\lambda}$, hence the equilibrium cannot be learned by least-squares methods.

Alternatively, the Eigenvalues of the matrix $G(\lambda^*)$ in the Incomplete Information case are given by the vector $(-\frac{1}{\theta}, 0, 0)$, where $|-1/\theta| < 1$ which, together with the continuity of $G(\lambda)$ gives necessity. Sufficiency comes by demonstrating E-stability of the learning algorithm. This is done by analyzing the limiting mapping $T(\lambda)$ defined as

$$T(\lambda) \equiv \lim_{t \rightarrow \infty} \mathbb{E}(H(\lambda, X_t(\lambda))) = \begin{pmatrix} S^{-1}\sigma_1 \\ \sigma_0 - S \end{pmatrix}, \quad (3.14)$$

where σ_i is the auto-covariance of order i of the forecast errors process e_t under λ . E-stability is verified by studying the stability of the differential equation

$$\frac{d\lambda}{d\tau} = T(\lambda) - \lambda \quad (3.15)$$

in the neighborhood of λ^* . It is possible to show that λ^* is a stationary point of the above ordinary differential equation and that such stationary point is stable. This last property is obtained by deriving the partial derivatives matrix of $T(\lambda)$ evaluated at λ^* , which is⁸

$$DT(\lambda^*) = \begin{pmatrix} -\frac{(\theta+\kappa)^2}{\theta} & 0 \\ \tau_{S\eta}(\theta) & -1 \end{pmatrix}, \quad (3.16)$$

where $\tau_{S\eta}(\theta)$ is a function whose form is not influential for the stability analysis. The Eigenvalues of $DT(\lambda^*)$ must be all negative for the stationary point to be stable. Note that this is verified under our maintained assumption $\theta > 1$. Hence, the incomplete information equilibrium λ^* can be

⁸See Appendix A for details

learned by the least-squares algorithm (3.4)-(3.6).

We have established the following theorem.

Theorem 1. *The Full Information rational expectations equilibrium $\hat{\lambda}$ cannot be learned by the least-squares algorithm (3.4)-(3.6) for any λ_0 in a neighborhood of $\hat{\lambda}$. The Incomplete Information equilibrium λ^* can always be learned by the least-squares algorithm (3.4)-(3.6) for any λ_0 in a neighborhood of λ^* .*

3.2 INFORMATIONAL (IN)STABILITY The reason behind the result of Theorem 1 lies in the structure of the dynamic process for the forecast errors e_t . As we demonstrate below, there is a hidden instability in the forecast errors associated with the Full Information equilibrium that is not present in the Incomplete Information case. This instability implies that any perturbation in the agents' information set will cause the Full Information equilibrium to become non-stationary. The learning algorithm exposes this instability.

To substantiate this point, let $a_t = u_t + \theta u_{t-1}$ with $|\theta| > 1$, as above, and assume that the representative agent is endowed with the knowledge of some state u_0 , so that $u_0 \in \mathbb{U}_t$ for any $t \geq 0$. Proposition 1 gives the Full Information equilibrium as

$$y_t = \varphi(1 + \theta\kappa)u_t + \varphi\theta u_{t-1}. \quad (3.17)$$

Note that (3.17) implies $\mathbb{E}[y_{t+1}|\mathbb{U}_t \vee \mathbb{V}_t(y)] = \varphi^{-1}\mathbb{E}[\theta u_t|\mathbb{U}_t \vee \mathbb{V}_t(y)]$. Using (2.3) lagged once one can write

$$\begin{aligned} \mathbb{E}[\theta u_t|\mathbb{U}_t \vee \mathbb{V}_t(y)] &= \mathbb{E}[\theta u_t + \theta^2 u_{t-1} - \theta^2 u_{t-1}|\mathbb{U}_t \vee \mathbb{V}_t(y)] \\ &= \theta a_t - \mathbb{E}[\theta^2 u_{t-1}|\mathbb{U}_t \vee \mathbb{V}_t(y)] \\ &= \theta a_t - \theta \frac{\varphi}{\kappa}(y_{t-1} - a_{t-1}). \end{aligned}$$

Once substituted into the equilibrium equation one obtains

$$y_t = -\theta y_{t-1} + \varphi \left((1 + \theta\kappa) + \theta L \right) a_t, \quad (3.18)$$

which must hold in the Full Information rational expectations equilibrium (3.17).

Note that in this representation, the equilibrium variable y_t has an autoregressive root at $|\theta| > 1$ that is explosive, in the sense that any realization of the innovation different from zero would result in non-stationary behavior of the variable y_t . Despite the unstable autoregressive root, however, y_t is still a stationary process, as equation (3.17) clearly shows. The reason that reconciles this apparent contradiction is that the autoregressive root of y_t in (3.18) exactly cancels with the moving average root of a_t : this potential instability is exactly defused by the dynamics of a_t . In the optimal control literature jargon, the representation (3.18) is said to harbor a “non-minimum phase zero” which can be canceled with the equivalent pole. Mathematically such cancellation is legitimate, but for the optimal control problem it creates a “hidden instability”: if the system is slightly

misspecified, the cancellation fails and the hidden pole drives the system to instability.⁹

To see how the instability can manifest, consider equation (3.18) at time $t = 1$,

$$y_1 = \varphi\theta a_0 - \theta y_0 + \varphi(1 + \theta\kappa)a_1. \quad (3.19)$$

The condition for unstable root to be exactly diffused by the root of the a_t process is that a_0 and y_0 must be specified as

$$\varphi a_0 - y_0 = \frac{\kappa}{\theta} u_0. \quad (3.20)$$

Because of the hidden instability, the dynamic equation (3.18) must be initialized at the initial condition (3.20) in order for it to be stationary. If the initialization is done at a different point, the unstable root would make the process explosive. To see this, suppose that instead of observing u_0 perfectly, the representative agent is given a noisy signal $v_0 = u_0 + \frac{\theta}{\kappa}\sigma_\varepsilon\varepsilon$, where $\sigma_\varepsilon > 0$ and $\varepsilon \sim N(0, 1)$. To evaluate stability we maintain that the agent acts under the assumption that $\sigma_\varepsilon = 0$. The initial condition is then given by

$$\varphi a_0 - y_0 = \frac{\kappa}{\theta} v_0. \quad (3.21)$$

Noting that $\frac{\kappa}{\theta}v_0 = \frac{\kappa}{\theta}u_0 + \sigma_\varepsilon\varepsilon$, and using (3.18) to solve for the dynamics forward one obtains

$$y_t = \varphi(1 + \theta\kappa)u_t + \varphi\theta u_{t-1} + \theta^t\sigma_\varepsilon\varepsilon, \quad (3.22)$$

which is non-stationary and diverging with respect to (3.17).

While we interpret $\sigma_\varepsilon > 0$ as a perturbation to the information set of the representative agent, one could also interpret the realization $\sigma_\varepsilon\varepsilon \neq 0$ as a perturbation to κ in (3.21), in which case the perturbation is to the knowledge of the structure of the model, as in a standard learning exercise. Solving the dynamics forward using (3.18) one would then also obtain (3.22), where the initial perturbation results in unstable dynamics. This shows how, in the context of our model, learning about realizations of shocks and learning about equilibrium parameters are closely related.

Conversely, the same decomposition for the Incomplete Information equilibrium yields

$$y_t = -\frac{1}{\theta}y_{t-1} + \frac{\varphi}{\theta}\left(\theta + \kappa + L\right)a_t \quad (3.23)$$

The equilibrium process y_t has an auto-regressive representation, as in (3.18), but the autoregressive root is now stationary, since $|\frac{1}{\theta}| < 1$. As opposed to the Full Information case, there are no hidden instabilities being created in the equilibrium dynamics. Consider equation (3.23) at time $t = 1$,

$$y_1 = \frac{1}{\theta}(\varphi a_0 - y_0) + \varphi\left(\frac{\theta + \kappa}{\theta}\right)a_1 \quad (3.24)$$

⁹See Skogestad and Postlethwaite (2005) for a textbook analysis of hidden instabilities in dynamic systems.

In this case, to ensure that y_1 is exactly equal to the equilibrium level after an arbitrary history of innovations u_t 's the initial condition should be specified as

$$\varphi a_0 - y_0 = \varphi \kappa \tilde{u}_0 \quad (3.25)$$

The second equality in this expression is intuitive: in the same way as the initial condition of the Full Information equilibrium was a linear function of u_0 , the initial condition for the Incomplete Information equilibrium is a linear function of \tilde{u}_0 , the innovation observed by the agent.

To evaluate the informational stability of (3.23) we proceed as we did for the Full Information case. Suppose that the agent's information set is perturbed by providing the agent with the noisy signal $w_0 = \tilde{u}_0 + \frac{1}{\varphi \kappa} \sigma_\varepsilon \varepsilon$, where $\sigma_\varepsilon > 0$ so that

$$\varphi a_0 - y_0 = \varphi \kappa w_0. \quad (3.26)$$

Since $\varphi \kappa w_0 = \tilde{u}_0 + \sigma_\varepsilon \varepsilon$, the dynamic solution to (3.23) becomes

$$y_t = \varphi(\theta + \kappa + L)\tilde{u}_t + \frac{1}{\theta^t} \varepsilon. \quad (3.27)$$

Under $|\frac{1}{\theta}| < 1$ the accumulated effect of the perturbation disappears as t grows larger and the perturbed equilibrium converges back to the original equilibrium. Hence, the Incomplete Information equilibrium is stable.

The following proposition generalizes the examples.

Proposition 2. *If $m > 0$ ($A(L)$ is non-invertible), the Full Information equilibrium contains a hidden instability. If the information set is perturbed ($\sigma_\varepsilon > 0$), the dynamics will become non-stationary and diverge away from the full information equilibrium. Conversely, the Incomplete Information equilibrium does not contain a hidden instability. If the information set is perturbed ($\sigma_\varepsilon > 0$), the dynamics of the perturbed system will revert back to the original dynamics.*

Proof. See Appendix A. □

This proposition provides the intuition behind the result of Theorem 1. The Full Information equilibrium contains an explosive root which exactly cancels with the zero in the moving average of a_t . If the forecast error process is not initialized at the equilibrium value, the explosive root will not cancel with the zero and the initial misalignment would grow indefinitely.

In the learning algorithm, the forecast error is not initialized at the equilibrium value and so the explosive root prevents the convergence of the learning algorithm. The Incomplete Information equilibrium does not contain a hidden instability. Even though the forecast error is not initialized at the equilibrium value, the learning algorithm converges and is E-stable.

To make this point more concrete, consider first the full information equilibrium. Suppose that the learning algorithm is initiated at the point $\eta_0 = \hat{\eta}$, where $\hat{\eta} = \frac{\theta}{1+\kappa\theta}$. In addition let $S_0 = \varphi^2(1 + \kappa\theta)^2$. Under these values, if one were to assume that the initial forecast error e_0 is

equal to what one would get in the RE equilibrium, i.e. $\varphi(1 + \theta\kappa)u_0$, then beliefs would remain at their correct initial value $\hat{\eta}$ through the entire algorithm. In accordance with our exercise for informational stability, we instead assume that the initial forecast error is perturbed by the disturbance ε , so that $e_0 = \varphi(1 + \theta\kappa)(u_0 + \varepsilon)$. Working through the algebra one can show that the beliefs evolve as $\eta_t = \hat{\eta} + \sigma_\varepsilon^2 \hat{c}_t(\theta)$, where $\hat{c}_t(\theta)$ is a ratio of polynomials in θ with t being the difference between the highest degree of the numerator and the highest degree of the denominator. For $t \rightarrow \infty$ one then has $\hat{c}_t \rightarrow \infty$ when $\theta > 1$, which shows how the beliefs under the learning algorithm diverge off the equilibrium value. It is instructive for the purpose of this discussion to consider the case of $t = 1$, which results in $\eta_t = \hat{\eta} - \sigma_\varepsilon^2 \theta$. In other words, the first deviation from the equilibrium value is a function of the variance of the disturbance ε , and the size of θ , so that the larger is θ , the larger the deviation.

For the Incomplete Information equilibrium we perform the same exercise and set $\eta_0 = \eta^*$, where $\eta^* = \frac{1}{\theta + \kappa}$, and let $S_0 = \varphi^2(\theta + \kappa)^2$. Once again, by setting e_0 to equal what one would get in the RE equilibrium, i.e. $\varphi(\kappa + \theta)\tilde{u}_0$, then beliefs would remain at their correct initial value η^* . We perturb the initial forecast error by setting $e_0 = \varphi(\kappa + \theta)(\tilde{u}_0 + \varepsilon)$. The pattern of beliefs for η can now be summarized by $\eta_t = \eta^* + \sigma_\varepsilon^2 c_t^*(\frac{1}{\theta})$, where $c_t^*(\frac{1}{\theta})$ is a ratio of polynomials in $1/\theta$ with t being the difference between the highest degree of the numerator and the highest degree of the denominator. In this case one cannot readily conclude that the limiting behavior rests on η^* , although we know, from the analysis of Section 3.1, it eventually does. However, by looking at the case of $t = 1$, one can see that the first step of the learning algorithm implies $\eta_t = \eta^* - \sigma_\varepsilon^2 \frac{1}{\theta}$, which clearly show that the larger is the size of θ , the smaller the deviation from the true equilibrium value.

As a final comment on the discussion above, we point out that the divergence of the learning algorithm in the Full Information case is due to the divergence of the forecast error process e_t . Unless that process is initialized at the exact equilibrium value, the learning algorithm has a built-in non-stationarity (which lies in the Eigenvalue of the matrix $G(\lambda)$), such that the forecast errors build up and eventually diverge.

4 APPLICATION: PRODUCTIVITY SHOCKS WITH NOMINAL RIGIDITY

In this section we apply our results to an incomplete information model with nominal rigidities and productivity shocks. The specific form of the model is taken from Lorenzoni (2009), and it represents an instance of the class of monetary models popularized by Clarida, Gali, and Gertler (1999) and Woodford (2003b), among others.

The linearized economy is fully described by an output equation, an inflation equation and a

monetary policy rule as

$$y_t = \mathbb{E}(y_{t+1}|\Omega_t) - i_t + \mathbb{E}(\pi_{t+1}|\Omega_t) \quad (4.1)$$

$$\pi_t = \alpha(y_t - a_t) + \beta\mathbb{E}(\pi_{t+1}|\Omega_t) \quad (4.2)$$

$$i_t = \phi\pi_t, \quad (4.3)$$

where all the constants have been dropped for convenience, and $\alpha > 0$, $\beta \in (0, 1)$ and $\phi > 0$. Ω_t represents the information set of the representative agent, to be specified shortly. The variable a_t represents the exogenous process for aggregate productivity in the economy, while the term $y_t - a_t$ is a measure of the real marginal costs along the labor market equilibrium. In Lorenzoni (2009) the productivity process a_t is assumed to be equal to the sum of a permanent component and a transitory component. Agents are interested in forecasting future productivity and so they engage in predicting the unobserved permanent component, but the transitory component acts as noise in their signal extraction problem. We follow Lorenzoni (2009) and specify a process for aggregate productivity with a permanent component modeled as a stochastic trend, but instead of superimposing a transitory component we allow for a richer moving average structure, which, as will become clearer, still prevents agents from disentangling the permanent component in productivity. More specifically we let

$$a_t = a_{t-1} + u_t + \theta u_{t-1}, \quad (4.4)$$

with $u_t \sim \mathcal{N}(0, \sigma_u^2)$. We assume that $\theta > 1$, which results in productivity a_t following a diffusion process with a typical S-shape behavior: for a given innovation u_t , productivity responds slowly at first - with an impact coefficient of 1, and more strongly subsequently - with a delayed impact coefficient of $\theta > 1$, before leveling-off to the new permanent level $(1+\theta)u_t$. Arguably, such dynamic behavior adequately captures episodes of innovations in productivity that have permanent effects (cf. Canova (2003)). Importantly, in our setting the slow diffusion dynamics also affects available information. Under (4.4), the observation of the current and past realizations of a_t is insufficient to recover the current and past realizations of u_t , which results in agents engaging in a dynamic signal extraction problem in order to predict a_{t+1} .

Our objective is to characterize the dynamic rational expectations equilibrium with incomplete information. To that end we assume, for simplicity, that the initial time is not specified, we let $\mathbb{U}_t = \{\emptyset\}$ for all t 's and define the information set as $\Omega_t \equiv \mathbb{V}_t(y, \pi)$. From the structural equations (4.1)-(4.3) agents are always able to compute aggregate productivity a_t , and so in any equilibrium it will be that $\mathbb{V}_t(a) \subseteq \mathbb{V}_t(y, \pi)$. We conjecture that $\mathbb{E}(y_{t+1}|\Omega_t) = \mathbb{E}(a_{t+1}|\Omega_t)$, which implies $\mathbb{E}(\pi_{t+1}|\Omega_t) = 0$.¹⁰ Substituting the interest rate rule into the output equation and the inflation

¹⁰To see this just lead the inflation equation forward one period and apply the expectational operator on both sides, taking into account the conjecture for the expectations of output and aggregate productivity. In formulating our conjecture we closely follow Lorenzoni (2009).

equation under the above conjecture the equilibrium condition reduces to

$$y_t = \kappa \mathbb{E}[y_{t+1} | \Omega_t] + (1 - \kappa)a_t, \quad (4.5)$$

where $\kappa \equiv \frac{1}{1+\alpha\varphi} \in (0, 1)$. The equilibrium equation we want to study is thus of the form (2.1) with $\varphi = (1 - \kappa)$. Once the solution for y_t is obtained, the solution for π_t immediately follows.

The Full Information equilibrium of (4.5) is a straightforward application of Case [.ii] in Proposition 1 and it is given by the $ARMA(1, 1)$ process

$$\tilde{y}_t - \tilde{y}_{t-1} = (1 + \theta\kappa)u_t + \theta(1 - \kappa)u_{t-1}. \quad (4.6)$$

To evaluate the informational stability properties of (4.6) we express the equilibrium equation in terms of observed variables

$$\tilde{y}_t - \tilde{y}_{t-1} = -\theta(\tilde{y}_{t-1} - \tilde{y}_{t-2}) + \theta(1 - \kappa)(a_t - a_{t-1}). \quad (4.7)$$

We know from Section 3.2 that the autoregressive root $-\theta$ dictates the informational stability of the equilibrium. If $\theta \leq 1$ the equilibrium is informationally stable, while for $\theta > 1$ the equilibrium is informationally unstable. In the latter case equation (4.7) has a hidden instability at θ , which cancels along the equilibrium path, but it affects the dynamics around the equilibrium if perturbed. Therefore, under our assumption that a_t follows an S-shaped diffusion pattern, i.e. under $\theta > 1$, focusing on equilibrium (4.6) would be a choice that is not robust to the information set initialization, as highlighted by Proposition 2, and thus not least-squares learnable by Theorem 1.

We next compute the informationally stable rational expectations equilibrium for (4.5). Application of Case [.iii] in Proposition 1 shows that the equilibrium is given by an $ARMA(2, 2)$ of the form

$$y_t^* - \frac{\theta - 1}{\theta}y_{t-1}^* - \frac{1}{\theta}y_{t-2}^* = \frac{\theta + \kappa}{\theta}u_t + \left(\frac{1 - \kappa}{\theta} + \theta + \kappa\right)u_{t-1} + (1 - \kappa)u_{t-2}. \quad (4.8)$$

The equilibrium process inherits the unit root of the productivity process, and an additional stable auto-regressive root in the output process at $\frac{1}{\theta}$. Note that in equation (4.8) there are no hidden instabilities, which means that any perturbation around the equilibrium path would eventually dissipate over time.

We have thus established the following corollary to Theorem 1.

Corollary 1. *In the economy with nominal rigidity and productivity shocks described by equations (4.1)-(4.4), the equilibrium (4.6) is not least-squares learnable, while equilibrium (4.8) is least-squares learnable.*

The corollary offers a criterion to select across the two equilibria, a choice that turns out to have important consequences in terms of the positive implications that the model can deliver. Figure 1 plots the impulse responses for productivity a_t (dotted line), the informationally unstable full in-

formation solution (4.6) (black dashed line), and the informationally stable incomplete information solution (4.8) (blue solid line) for both output and employment, to a one-time standard deviation increase in u_t . The parameters values are set to $\alpha = .05$, $\beta = .7$, $\phi = 1.5$, $\rho = 1$ and $\theta = 2$. Looking at the left-side panel, in the unstable solution output increases at impact close to the new permanent level, as the extent of the innovation is fully anticipated. The increase in productivity is limited at impact because of the initial slow diffusion, which means that employment has to increase in order for output to be higher at impact, which can be seen in the right-side panel. In the subsequent period productivity reaches the new permanent level, so does output, and employment goes back to its long run trend. In the unstable solution, the economy's response to a supply shock has a profile that resembles a response to a demand shock, but it only lasts one period.

Consider now the unstable solution. The initial reaction of output is positive, but is around only 50% of the new higher trend. This level is still higher than the increase in productivity and so employment increases at impact, but only by a 25% of the increase in the unstable case. One period from impact agents learn that the permanent increase in productivity might actually be higher than previously thought, and output overshoots the trend by around a 20% margin. Productivity is now at the new higher level, but the overshooting in output requires a higher employment to be achieved, which means that employment is still above trend one period after impact, and - in our numerical example - still higher than the level at impact. The subsequent period agents realize that the innovation might have been smaller than they thought, which creates a drop in output below trend, and a drop in employment below trend. In other words, two periods after the onset of the productivity innovation and a gradual boom in output and employment, the economy experiences a recession. The same type of oscillation then repeats itself with declining magnitude, until the economy eventually settles on the new higher trend.

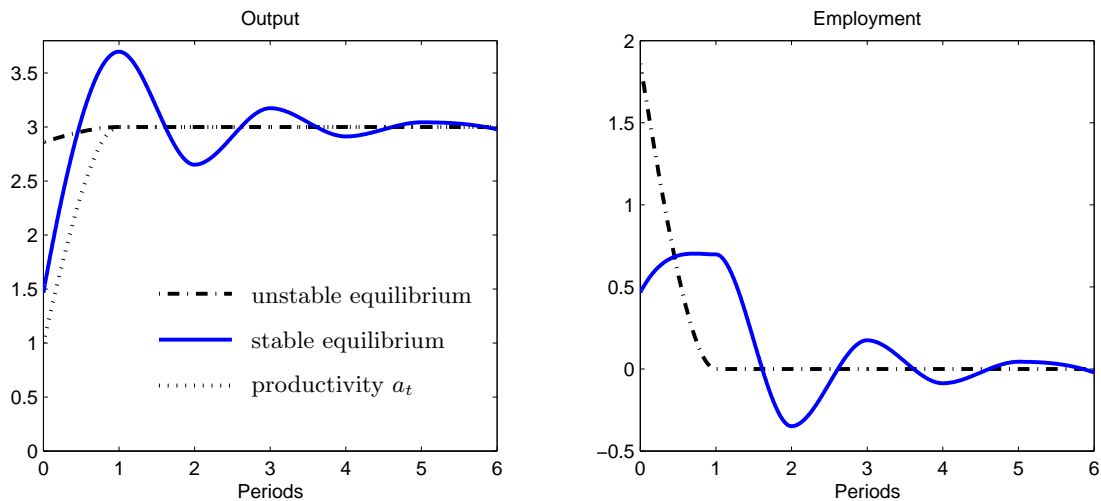


Figure 1: Impulse Response of Output and Employment to Innovation in Productivity

In summary, applying our analysis to the equilibrium of the stylized monetary model (4.1)-

(4.3) when the productivity process a_t displays a diffusion-type dynamics, results in a rational expectations equilibrium with several differences compared to the equilibrium on which a researcher unaware of our analysis may focus. From a normative perspective, the equilibrium (4.8) is robust to perturbations in the information set and is least-squares learnable, two conditions that would make it preferable to equilibrium (4.6). From a positive perspective, the equilibrium (4.8) displays a rich propagation dynamic that not only generates a demand-shock like response at impact, which is qualitatively similar to what happens in equilibrium (4.6), but also generates a longer expansion of output and employment over trend, and, remarkably, a subsequent recession in both output and employment (both below trend), followed by a smaller expansion, and so on. This all in response to a permanent positive productivity innovation with a one-period gradual diffusion in the economy.

5 CONCLUSION

In this paper, we studied the least-squares learnability of Rational Expectations equilibria in dynamic models with incomplete information. We showed that equilibria where the endogenous variables resolve the information incompleteness cannot be learned and possess hidden instabilities, in the sense that a slight perturbation in the endogenous information set of the agents along the equilibrium path can lead to explosive dynamics. We then presented a class of dynamic rational expectations equilibria that are learnable for the same parameter space. We concluded by presenting an application from an established literature.

Given the importance of learnability in models of rational expectations equilibrium, understanding the extent to which our results generalize would be of interest. In particular, analyzing the relationship between informational stability and learnability in a broader class of models would be a useful extension and is left to future research. Another obvious future line of work would be to estimate models with incomplete information. Learnable equilibria, it has been argued, are more likely to occur in reality [Evans and Honkapohja (2001)]. Examining the empirical fit of models with the particular form of non-revealing equilibria discussed here would be an interesting extension.

6 APPENDIX A: PROOFS

PROOF OF PROPOSITION 1 The Information Equilibrium will be a Full Information Equilibrium if the Hilbert space generated by Ω_t spans the space of the sequence of underlying structural shocks, $\{u_{t-j}\}_{j=0}^{\infty}$. The first two cases are straightforward: [i.] If $m = 0$, $\{a_{t-j}\}_{j=0}^{\infty}$ spans the same space as $\{u_{t-j}\}_{j=0}^{\infty}$ because a_t is fundamental for u_t . Therefore even if the exogenous information \mathbb{U}_t is the empty set, the Information Equilibrium will be the Full Information Equilibrium as a_t can always be inferred from the structural relationship (2.1). [ii.] If $m > 0$ and $\{u_{t-j}, u_{t-j-1}, \dots, u_{t-j-m+1}\} \in \mathbb{U}_t$ with $0 < j < \infty$, then by initializing the state at some time $t < 0$, agents have access to the full knowledge of the innovations u_t from t onward and the Information Equilibrium will be a Full Information Equilibrium. As an example, suppose a moving average representation has two distinct zeros, $x_t = (L - \lambda_1)(L - \lambda_2)\varepsilon_t$ with $|\lambda_1| < 1$, $|\lambda_2| < 1$. If agents observe ε_{t-1} and ε_{t-2} directly, then we may write the moving average as $x_t - \varepsilon_{t-2} + (\lambda_1 + \lambda_2)\varepsilon_{t-1} = \lambda_1\lambda_2\varepsilon_t$, which is always invertible. When $j = \infty$, technically speaking the initialization of the state is not well defined and so the condition on the exogenous information set \mathbb{U}_t has no bite, and the existence of a Full Information equilibrium cannot be excluded.

For case [iii.], we begin by guessing a functional form for the equilibrium price as

$$y_t = Q(L) \prod_{i=1}^m (1 - \lambda_i z) \tilde{u}_t \quad (6.1)$$

with

$$\tilde{u}_t = \prod_{i=1}^m \mathcal{B}_{\lambda_i}(L). \quad (6.2)$$

and

$$\mathcal{B}_{\lambda_i}(L) \equiv \frac{1 - \lambda_i L}{\lambda_i - L}. \quad (6.3)$$

Under such guess one can derive the conditional expectation for future productivity and substitute it into the equilibrium equation (2.1) and get the following z -transform expression

$$\begin{aligned} Q(z) \prod_{i=1}^m (z - \lambda_i) &= \kappa z^{-1} [Q(z) \prod_{i=1}^m (1 - \lambda_i z) - Q_0] \prod_{i=1}^m \mathcal{B}_{\lambda_i}(z) + \varphi A(z) \\ &= \kappa z^{-1} [Q(z) \prod_{i=1}^m (z - \lambda_i) - Q_0 \prod_{i=1}^m \mathcal{B}_{\lambda_i}(z)] + \varphi A(z) \end{aligned}$$

Working out the algebra yields

$$Q(z)(z - \kappa) \prod_{i=1}^m (z - \lambda_i) = \varphi z A(z) - Q_0 \prod_{i=1}^m \mathcal{B}_{\lambda_i}(z) \quad (6.4)$$

For $|\kappa| < 1$, stationarity requires the $Q(\cdot)$ process to be analytic inside the unit circle, which will not be the case unless the process vanishes at the poles $z = \{\lambda_i, \kappa\}$ for every i . For simplicity, we

assume $\lambda_i \neq \lambda_j$ for any $i \neq j$, however this restriction can be relaxed [see, Whiteman (1983)]. Evaluating at $z = \lambda_i$ provides a restriction on the $A(\cdot)$ process,

$$A(\lambda_i) = 0 \quad \text{for } i = 1, \dots, m, \quad (6.5)$$

which implies that $\lambda_i = -1/\theta_i$ for all i . By Proposition 10.4 of Conway (1991), this restriction guarantees that the knowledge of the model does not reveal any additional information than the posited price sequence. Finally, evaluating (6.4) at $z = \kappa$ gives

$$Q_0 = \frac{\kappa A(\kappa)}{\prod_{i=1}^m \mathcal{B}_{\lambda_i}(\kappa)} \quad (6.6)$$

Substituting this into (6.4) and rearranging the algebra returns expression (2.10).

PROOF OF PROPOSITION 2

The proof of the proposition consists in expressing the Full Information equilibrium in terms of the information set of the rational expectations equilibrium, similarly to (3.18) and then argue that the expression is unstable with respect to the initialization of the information set. The Full Information equilibrium for $m > 0$ can be expressed as

$$(y_t - \rho y_{t-1}) \prod_{i=1}^m (1 + \theta_i L) = \varphi \frac{\left(LA(L)(1 - \rho\kappa) - \kappa A(\kappa)(1 - \rho L) \right)}{(L - \kappa)(1 - \rho\kappa)\tilde{A}(L)} (a_t - \rho a_{t-1}). \quad (6.7)$$

The lag polynomial on the right hand side has a zero at κ by construction that cancels with the root at the denominator. It follows that the polynomial is stationary. On the other hand, the equilibrium representation for y_t has an $AR(m+1)$ component, where m of the roots are inside the unit circle. Unless the initial conditions on the information set $(y_0, y_{-1}, \dots, y_{-m})$ and $(a_0, a_{-1}, \dots, a_{-m})$ are chosen to exactly cancel the explosive roots, the equilibrium dynamics will diverge.

PROOF OF THEOREM 1

To complete the proof of Theorem 1 we need to evaluate the stability of the mapping

$$T \begin{pmatrix} \eta \\ S \end{pmatrix} = \begin{pmatrix} T_\eta(\eta, S) \\ T_S(\eta, S) \end{pmatrix} = \begin{pmatrix} S^{-1}\sigma_1(\eta) \\ \sigma_0(\eta) - S \end{pmatrix} \quad (6.8)$$

where the two covariances are given by

$$\sigma_0(\eta) = \frac{\varphi^2}{(1 - \eta\kappa)^2 - \eta^2} \left(1 - \frac{2\eta}{1 - \eta\kappa} \theta + \theta^2 \right) \quad (6.9)$$

$$\sigma_1(\eta) = \frac{\varphi^2}{(1 - \eta\kappa)^2 - \eta^2} \left(\theta - \frac{\eta}{1 - \eta\kappa} \right) \left(1 - \frac{\eta\theta}{1 - \eta\kappa} \right) \quad (6.10)$$

We use these expressions to compute the matrix of partial derivatives

$$DT \begin{pmatrix} \eta \\ S \end{pmatrix} = \begin{pmatrix} DT_\eta(\eta, S) \\ DT_S(\eta, S) \end{pmatrix} = \begin{pmatrix} \frac{dT_\eta}{d\eta}(\eta, S) & \frac{dT_\eta}{dS}(\eta, S) \\ \frac{dT_S}{d\eta}(\eta, S) & \frac{dT_S}{dS}(\eta, S) \end{pmatrix} \quad (6.11)$$

evaluated at $\eta = \eta^* = \frac{1}{\theta + \kappa}$ and $S = S^* = \varphi^2(\theta + \kappa)^2$. Proceeding with the algebra one obtains matrix (3.16), and the least-squares learning convergence immediately follows.

7 APPENDIX B: MULTIVARIATE EXTENSION

The results of Theorem 1 and Proposition 2 extend beyond univariate settings. In this section we describe how to proceed in order to generalize those results to multivariate settings. Consider the generic multivariate rational expectations model

$$\Gamma_0 y_t = \Gamma_1 y_{t-1} + \Psi z_t + \Pi \eta_t, \quad (7.1)$$

where y_t is an $n \times 1$ vector of endogenous variables, z_t is an $m \times 1$ vector of exogenous random shocks, η is a $k \times 1$ vector of expectation errors, which satisfy $E_t \eta_{t+1} = 0$ for all t . Γ_0 and Γ_1 are $n \times n$ coefficient matrices, along with Ψ ($n \times m$) and Π ($n \times k$). The model collapses to the univariate setting of Section 2 when $z_t = u_t + \theta u_{t-1}$.

Sims (1998) uses a generalized Schur decomposition of Γ_0 and Γ_1 to show that there exist matrices such that $Q' \Lambda Z' = \Gamma_0$, $Q' \Omega Z' = \Gamma_1$, $Q' Q = Z' Z = I_{n \times n}$, where Λ and Ω are upper-triangular. The ratios of the diagonal elements of Ω and Λ , ω_{ii}/λ_{ii} , are the generalized eigenvalues. Defining $w_t = Z' y_t$ and pre-multiplying (7.1) by Q , yields the decomposition

$$\begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ 0 & \Lambda_{22} \end{bmatrix} \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ 0 & \Omega_{22} \end{bmatrix} \begin{bmatrix} w_{1,t-1} \\ w_{2,t-1} \end{bmatrix} + \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} (\Psi z_t + \Pi \eta_t) \quad (7.2)$$

The system is partitioned so that the generalized eigenvalues imply an explosive path for $w_{2,t}$ and a stable path for $w_{1,t}$. To ensure stability of the system, $w_{2,t}$ must be solved forward. Sims shows that the forward solution of (7.1) is

$$y_t = \Theta_1 y_{t-1} + \Theta_0 z_t + \Theta_y \sum_{s=1}^{\infty} \Theta_f^{s-1} \Theta_z E_t z_{t+s} \quad (7.3)$$

where $H = Z \begin{bmatrix} \Lambda_{11}^{-1} & -\Lambda_{11}^{-1}(\Lambda_{12} - \Phi \Lambda_{22}) \\ 0 & I \end{bmatrix}$, $\Theta_0 = H \begin{bmatrix} Q_1 - \Phi Q_2 \\ 0 \end{bmatrix} \Psi$

and $\Theta_y = -H_2$, $\Theta_1 = Z_1 \Lambda_{11}^{-1} [\Omega_{11} (\Omega_{12} - \Theta \Omega_{22})] Z$, $\Theta_f = \Omega_{22}^{-1} \Lambda_{22}$, and $\Theta_z = \Omega_{22}^{-1} Q_2 \Psi$.¹¹

The most basic informational assumption that will deliver non-invertibility is $z_t = \epsilon_{i,t-q}$ for some i , which is non-invertible because the moving average has a zero inside the unit circle at

¹¹We assume that the conditions necessary for a unique solution to exist hold. Specifically that the row space of $Q_1 \Pi$ be contained in that of $Q_2 \Pi$ [See Sims (1998)].

$L = 0$. If the agents do not observe the structural shocks, ϵ_{it} , (i.e., $\mathbb{U}_t = \{0\}$ for $t \in \mathbb{Z}$) then the last term in (7.3) drops out of the solution. Under this information structure, the solution (7.3) emits a stable vector-autoregression representation in current and past observables. Thus there are no hidden instabilities in the model.

However, if the agents observe the structural shocks directly (i.e., $\mathbb{U}_t = \{\epsilon_{i,t-j}\}_{j=0}^{\infty}$ for $i = 1, \dots, m$), the equilibrium is given by

$$y_t = \Theta_1 y_{t-1} + \Theta_0 \epsilon_{t-q} + \Theta_y \Theta_z [\epsilon_{t-q+1} + \Theta_f \epsilon_{t-q+2} + \dots + \Theta_f^{q-1} \epsilon_t] \quad (7.4)$$

which is the multivariate analog of (3.17). The term Θ_f is the multivariate analog to θ^{-1} in Section 3. In order to apply Proposition 2, we must show that the equilibrium is non-invertible in current and past y_t . Writing the equilibrium as a moving average, $y_t = \mathcal{A}(L)\epsilon_t$, a sufficient condition for non-invertibility is for $\det \mathcal{A}(L)$ to have a zero inside the unit circle. This will, of course, depend upon the model itself and parameterization of the model.

REFERENCES

- ADAM, K. (2007): “Optimal Monetary Policy with Imperfect Common Knowledge,” *Journal of Monetary Economics*, 54(2), 267–301.
- ALLEN, F., S. MORRIS, AND H. SHIN (2006): “Beauty Contests and Iterated Expectations in Asset Markets,” *Review of Financial Studies*, 19(3), 719–752.
- ANGELETOS, G., AND J. LA’O (2009): “Noisy Business Cycles,” *NBER Macroeconomics Annual*, 24.
- ANGELETOS, G., AND A. PAVAN (2007): “Efficient Use of Information and Social Value of Information,” *Econometrica*, 75(4).
- ANGELETOS, G. M., AND J. LA’O (2011): “Decentralization, Communication, and the Origins of Fluctuations,” Discussion Paper 17060, National Bureau of Economic Research.
- BACCHETTA, P., AND E. VAN WINCOOP (2006): “Can Information Heterogeneity Explain the Exchange Rate Puzzle?,” *American Economic Review*, 96(3), 552–576.
- CANOVA, F. (2003): *Methods for Applied Macroeconomic Research*. Princeton University Press, Princeton, New Jersey, first edn.
- CLARIDA, R., J. GALI, AND M. GERTLER (1999): “The Science of Monetary Policy: A New Keynesian Perspective,” *Journal of Economic Literature*, 37(4), 1661–1707.
- CONWAY, J. (1991): *The Theory of Subnormal Operators*. American Mathematical Society.
- EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton and Oxford.

- FUTIA, C. A. (1981): “Rational Expectations in Stationary Linear Models,” *Econometrica*, 49(1), 171–192.
- GRAHAM, L., AND S. WRIGHT (2010): “Information, Heterogeneity and Market Incompleteness,” *Journal of Monetary Economics*, 57(2), 164–174.
- GREGOIR, S., AND P. WEILL (2007): “Restricted Perception Equilibria and Rational Expectation Equilibrium,” *Journal of Economic Dynamics and Control*, 31(1), 81–109.
- HANSEN, L. P., AND T. J. SARGENT (1991): “Two Difficulties in Interpreting Vector Autoregressions,” in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent. Westview Press.
- HASSAN, T. A., AND T. M. MERTENS (2011): “The Social Cost of Near-Rational Investment,” NBER Working Paper 17027.
- HELLWIG, C. (2002): “Public Announcements, Adjustment Delays, and the Business Cycle,” mimeo, UCLA.
- HELLWIG, C. (2006): “Monetary Business Cycle Models: Imperfect Information,” *New Palgrave Dictionary of Economics*.
- HELLWIG, C., AND V. VENKATESWARAN (2009): “Setting the Right Prices for the Wrong Reasons,” *Journal of Monetary Economics*, 56, S57–S77.
- KASA, K. (2000): “Forecasting the Forecasts of Others in the Frequency Domain,” *Review of Economic Dynamics*, 3, 726–756.
- KASA, K., T. B. WALKER, AND C. H. WHITEMAN (2011): “Heterogenous Beliefs and Tests of Present Value Models,” University of Iowa Working Paper.
- LIPPI, M., AND L. REICHLIN (1994): “VAR Analysis, Nonfundamental Representations, Blaschke Matrices,” *Journal of Econometrics*, 63, 307–325.
- LORENZONI, G. (2009): “A Theory of Demand Shocks,” *American Economic Review*, 99(5), 2050–2084.
- NIMARK, K. (2011): “Dynamic Higher Order Expectations,” Universitat Pompeu Fabra Working Paper.
- PEARLMAN, J. G., AND T. J. SARGENT (2005): “Knowing the Forecasts of Others,” *Review of Economic Dynamics*, 8(2), 480–497.
- RONDINA, G. (2009): “Incomplete Information and Informative Pricing,” Working Paper. UCSD.
- SKOGESTAD, S., AND I. POSTLETHWAITE (2005): *Multivariable Feedback Control: Analysis and Design*. John Wiley and Sons, New York, 2 edn.

- TOWNSEND, R. M. (1983): “Forecasting the Forecasts of Others,” *Journal of Political Economy*, 91, 546–588.
- WHITEMAN, C. (1983): *Linear Rational Expectations Models: A User’s Guide*. University of Minnesota Press, Minneapolis.
- WOODFORD, M. (2003a): “Imperfect Common Knowledge and the Effects of Monetary Policy,” in *Knowledge, Information, and Expectations in Modern Macroeconomics*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford. Princeton University Press, Princeton, N.J.
- WOODFORD, M. (2003b): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, New Jersey.