Confounding Dynamics* 

Giacomo Rondina† Todd B. Walker‡ 

September 2017 

Abstract 

In the context of a dynamic model with incomplete information, we isolate a novel mechanism of shock propagation that results in waves of optimism and pessimism along a Rational Expectations equilibrium. We term the mechanism confounding dynamics because it arises from agents’ optimal signal extraction efforts on variables whose dynamics—as opposed to superimposed noise—prevents full revelation of information. Employing methods in the space of analytic functions, we are able to obtain analytical characterizations of the equilibria that generalize the celebrated Hansen-Sargent optimal prediction formula. We apply our results to a canonical one-sector real business cycle model. We show that, in response to a persistent positive productivity shock, confounding dynamics generate expansions and recessions that would not be present under complete information.

Keywords: Dispersed Information, Confounding Dynamics, Rational Expectations 

*We would like to thank Manuel Amador, Marios Angeletos, Ryan Chahrour, Tim Cogley, Thorsten Drautzburg, Jesús Fernández-Villaverde, Zhen Huo, Jennifer La’O, Eric Leeper, Thomas Mertens, Kristoffer Nimark, Gary Ramey, Tom Sargent, Martin Schneider, Dimitri Vayanos, Venky Venkateswaran, Pierre-Olivier Weill, Mirko Wiederholt, Chuck Whiteman and seminar participants at UC–Berkeley, UC–Santa Cruz, UC–San Diego, New York University, University of Pennsylvania, University of Texas–Austin, Stanford University, St. Louis Federal Reserve, Richmond Federal Reserve, Carnegie Mellon University, the 2017 Society of Economic Dynamics Annual Meeting in Edinburgh, and the 2017 Society for Computational Economics International Conference in New York. We also acknowledge financial support from the National Science Foundation under grant number SES-0962221. 

†University of California – San Diego, grondina@ucsd.edu 

‡Indiana University, walkertb@indiana.edu
1 INTRODUCTION

Modeling and seeking to understand economic fluctuations is one of the cornerstones of modern economics. The role of incomplete information in this endeavor was acknowledged very early on by Pigou (1929) and Keynes (1936). Their ideas were first formalized in a rational expectations setting by Lucas (1972, 1975), King (1982) and Townsend (1983b). The underlying theme that ties these papers together is that unresolved uncertainty—in and of itself—can be a source of fluctuation in the economy. This idea has seen a resurgence. Dynamic models with dispersed information are becoming increasingly prominent in several literatures such as asset pricing, optimal policy communication, international finance, and business cycles.\(^1\) Our paper contributes to this literature by introducing the concept of confounding dynamics and doing so in a manner that permits tractability.

Confounding dynamics arise from optimal prediction (i.e. rational expectations) in which past realizations of economic shocks prevent full revelation of information today, even when an arbitrarily large amount of data is available. Ensuring confounding dynamics survive amounts to deriving (non-invertibility) restrictions on equilibrium dynamics. If endogenous variables are non-invertible in current and past observations, agents will never fully unravel the contemporaneous economic shock. This is true even when the number of shocks is equal to the number of observables. Thus, a distinguishing feature of our approach is that confounding dynamics does not rely on the need to impose exogenous noise on the model.\(^2\)

The defining property of confounding dynamics is that the equilibrium impulse response functions consist of oscillating over- and under-reactions relative to the full information (or exogenous noise) equilibrium. The typical shape of an impulse response to a fundamental shock under complete information is one that monotonically decays back to steady state either from above or below. With confounding dynamics, impulse response functions fluctuate around the full information counterpart. That is, they display the “waves of optimism and pessimism” of Pigou (1929).

We articulate our idea in three steps. We first derive an optimal prediction formula under confounding dynamics that extends the celebrated Hansen-Sargent formula, and makes an explicit connection to these dynamics (see Section 2.2). Subsequently, we demonstrate that this behavior carries over to a generic rational expectations model with dispersed information. Our main theorem contains two equations—one that characterizes the dynamic properties of the equilibrium when confounding dynamics are present and one that derives exogenous restrictions that guarantee

---


\(^2\)Many of the papers in the dynamic dispersed information literature introduce exogenous noise in order to preserve asymmetric information in equilibrium. In Section 4.3 we show that our approach can accommodate exogenous noise, while still maintaining confounding dynamics.
confounding dynamics are preserved in equilibrium (see Section 4.2). Finally, we provide economic intuition by introducing confounding dynamics into a standard Real Business Cycle model. We find that the elasticity of intertemporal substitution in agents’ preferences plays a critical role both for the existence of confounding dynamics, and the degree by which they affect equilibrium allocations. To the best of our knowledge, this equilibrium dynamic behavior and formal characterization is new to the rational expectations literature (see Section 5).

We solve and analyze the rational expectations equilibrium in the space of analytic functions. This approach has several advantages vis-a-vis standard time-domain methods. For example, as emphasized in Townsend (1983a), equilibria are sought in generic functional spaces spanned by linear combinations of shocks, which allows one to avoid explicitly modeling higher-order belief dynamics (Section 3.5). Moreover, the matrix Ricatti equation is replaced by a more transparent spectral factorization problem. This allows us to solve and analyze the equilibrium in closed form. We are not the first to advocate such an approach. Others, such as Futia (1981), Townsend (1983a), Taub (1989), Kasa (2000), Walker (2007), Rondina (2009), Bernhardt, Seiler, and Taub (2010), Kasa, Walker, and Whiteman (2014), and Huo and Takayama (2016) have used similar techniques to solve dynamic rational expectation models with incomplete information. We contribute to this literature by deriving analytical representations (e.g., generalized Hansen-Sargent formulas) and by providing a systematic treatment of equilibrium conditions in models with dispersed information that display confounding dynamics. Futia (1981) and Townsend (1983a) were the first to advocate for the use of analytic functions to solve dynamic rational expectations models with heterogeneous information. Many of the mathematical antecedents of this paper can be found there and in Whiteman (1983). Taub (1989) demonstrates how the algebra associated with dynamic signal extraction (i.e., spectral factorization) is simplified through the analytic function approach. We take advantage of these formulas to completely characterize existence and uniqueness of equilibria in dispersed informational setups. Bernhardt, Seiler, and Taub (2010) and Kasa, Walker, and Whiteman (2014) do not examine models with dispersed information, but show how these methods can be used to help resolve asset pricing anomalies.

2 Preliminaries

In this section, we first establish notation and introduce relevant mathematical definitions. We then formalize the notion of confounding dynamics, presenting a simple example that shows the mechanism at work.

2.1 Mathematical Preliminaries Throughout the paper, we work in the space of polynomials in the lag operator $L$ with square-summable coefficients that operate on Gaussian random variables. In our framework, any stochastic process $\omega_t$ can always be written as

$$\omega_t = Q(L)\varepsilon_t = \sum_{j=0}^{\infty} Q_j L^j \varepsilon_t,$$

(2.1)
where $\sum_{j=0}^{\infty} |Q_j|^2 < \infty$, and $\varepsilon_t \sim N(0, \sigma_\varepsilon)$, are innovations identically and independently distributed over time. In linear-Gaussian environments, working with representations of the form of (2.1), and their functional equivalents, has three advantages for analyzing rational expectations models with incomplete information.$^3$

First, representation (2.1) is general in the sense that it can accommodate both auto-regressive (AR) and moving-average (MA) components of any order. This is especially useful when searching for an equilibrium because it avoids the need to specify a conjecture with a specific ARMA order. Regardless of the complexity of the equilibrium conditions that emerge in models of dispersed information (e.g., infinite regress in expectations), the solution will take the form of (2.1).$^4$

Second, the Wold Representation Theorem ensures that processes like $\omega_t$ can always be written uniquely as a linear combination of a moving average representation where the innovations are the linear forecast errors for $\omega_t$, conditional on any linear-Gaussian information set [Brockwell and Davis (1987)]. That is, the Wold representation establishes the invertibility of $Q(L)$ and one may write $Q(L)^{-1} \omega_t = \varepsilon_t$, which implies that the space spanned by $\{\omega_t, \omega_{t-1}, \ldots\}$ is equivalent (in mean-square norm) to the space spanned by $\{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$. Consequently, one can apply the optimal prediction formulas derived by Wiener-Kolmogorov [Whittle (1983)] to compute the conditional expectation of processes like $\omega_t$.

Third, the Riesz-Fischer Theorem [see Sargent (1987)] establishes an isometric, isomorphic mapping from the space of lag polynomials with square-summable coefficients $Q(L)$ to the space of analytic complex-valued functions, where (2.1) is represented as $Q(z)$, but with $z \in \mathbb{C}$. In several key steps of the analysis in this paper we find it convenient to exploit the properties of such functions, which allows us to derive simple existence and uniqueness conditions for rational expectations equilibria with incomplete and dispersed information following Whiteman (1983). In a slight abuse of notation, we employ $L$ and $z$ interchangeably when working in the space of analytic functions.

While this methodology is extremely helpful in solving dynamic models with incomplete information, it is not well known by economists. Thus, we provide Online Appendix C with the statements of the key theorems cited above and further references for interested readers. We now restrict our focus to the formulation of optimal prediction formulas, which is where confounding dynamics emerge.

Suppose that we would like to formulate the prediction of $\omega_{t+j}$ so as to minimize the mean-squared forecast error, conditional on the observation of the history of a $n \times 1$ vector of variables, $s_t$, up to time $t$. To denote such history, we use the compact notation, $s_t = \Gamma(L)u_t$, (2.2)
where \( \Gamma(L) \) is an \( n \times m \) matrix with each element being a square-summable lag polynomial in non-negative powers of \( L \), and \( u_t \) an \( m \times 1 \) vector of i.i.d. Gaussian shocks with variance-covariance normalized to the \( m \times m \) identity matrix. Let \( g_{ss}(z) \) be the variance-covariance generating function for the process \( s_t \), then one has that \( g_{ss}(z) = \Gamma(z)\Gamma(z^{-1})^\top \). Similarly, one can define the covariance generating function between the joint processes \( \omega_t \) and \( s_t \), which is given by \( g_{ws}(z) = Q(z)\sigma_e\Gamma(z^{-1})^\top \). The prediction for \( \omega_{t+j} \) that minimizes the mean squared forecast error corresponds to a linear combination of current and past realizations of \( s_t \), denoted by

\[
\mathcal{P}(\omega_{t+j}|s^t) = \Pi(L)s_t.
\]

Here, \( \Pi(L) \) is a \( 1 \times n \) vector of square-summable lag polynomials in non-negative powers of \( L \), whose form is provided by the Wiener-Kolmogorov formula:

\[
\Pi(L) = \left[ L^{-j}g_{ws}(L)(\Gamma^*(L^{-1})^\top)^{-1}\right]_+\Gamma^*(L)^{-1}.
\]

Expression (2.4) has several moving parts that require some unpacking. Let us first consider the special case \( s_t = \omega_t \), which implies \( \Gamma(L) = Q(L) \) and \( u_t = \varepsilon_t \), so that the prediction problem is one in which we would like to predict the future realizations \( \omega_{t+j} \), for \( j \geq 1 \), using its own past, \( \omega^j \). Given the form of \( \omega_t \) in (2.1), the best prediction is one that carries \( Q(L) \), \( j \) periods forward and uses the best estimates of the infinite history of innovations \( \{\varepsilon_{t+j},\varepsilon_{t+j-1},...\} \), to compute \( \omega_{t+j} \). The best estimates of \( \{\varepsilon_{t+j},\varepsilon_{t+j-1},...\} \) are clearly equal to the unconditional average, 0. It follows that they should not appear in the optimal prediction. This is achieved by the operator \([ ]_+\) in (2.4), known as the “annihilating operator”, which instructs us to ignore the first \( j \) coefficients of its argument.

The best estimates for \( \{\varepsilon_t,\varepsilon_{t-1},\varepsilon_{t-2},...\} \), on the other hand, are usually not zero, and correspond to the innovations in the Wold fundamental representation for \( \omega_t \). The decomposition of the information set into its Wold fundamental representation is achieved by \( \Gamma^*(L) \) in (2.4). Whether \( \Gamma^*(L) \) is equal to \( \Gamma(L) \) depends on the invertibility properties of the analytic function \( \Gamma(z) \). For our special case, if \( Q(z) \) is invertible for \( z \) inside the unit circle, then the Wold fundamental representation corresponds exactly with the history \( \{\varepsilon_t,\varepsilon_{t-1},\varepsilon_{t-2},...\} \), and the prediction formula \( \mathcal{P}(\omega_{t+j}|\omega^j) = \tilde{\Pi}(L)\omega_t \), has

\[
\tilde{\Pi}(L) = [L^{-j}Q(L)]_+Q(L)^{-1}.
\]

Note that all the steps just described are clearly at work here: \( L^{-j} \) carries the function \( Q(L) \) \( j \) periods forward; the operator \([ ]_+\) sets the estimates of innovations from \( t+1 \) to \( t+j \) to zero, by annihilating \( Q_0, Q_1,...,Q_{j-1} \); finally, \( Q(L)^{-1} \) makes sure that once \( \Pi(L) \) is multiplied by \( \omega_t \), the Wold fundamental innovations \( \{\varepsilon_t,\varepsilon_{t-1},\varepsilon_{t-2},...\} \) result.

5The variance-covariance generating function of a stationary Gaussian process is defined as the Fourier transform of its correlogram, which is the collection of covariances at all horizons. See the Online Appendix C for details.

6See Whittle (1983), or the Online Appendix C, for a derivation of the formula.
Expression (2.5) is a special case of (2.4) because, in general, $\Gamma(z)$ is not invertible for $z$ inside the unit circle, and solving for $\Gamma^*(z)$ is at the core of the solution to the prediction problem. Formally, $\Gamma^*(z)$ corresponds to the “canonical” factorization of the variance-covariance generating function $g_{ss}(z)$, so that

$$g_{ss}(z) = \Gamma^*(z)\Gamma^*(z^{-1})^T. \quad (2.6)$$

The canonical factorization answers the question: What space is spanned by the observables $s^t$? If $\Gamma(z)$ is invertible in non-negative powers of $z$, then $s^t$ will span the innovations $u^t$. If $\Gamma(z)$ is non-invertible, then $\Gamma^*(z)$ must be determined and it will span a space that is strictly smaller than that spanned by $u^t$. The existence of $\Gamma^*(z)$ is guaranteed by the Wold (fundamental) Representation Theorem. The computation of $\Gamma^*(z)$ can be quite involved, however, and there exist several methods for achieving it. In our setting, we follow the steps outlined in Roanov (1967).

The most straightforward reason for $\Gamma(z)$ to be non-invertible for prediction purposes, is when $m > n$. In such case, the dimension of the vector of underlying shocks, $u_t$, is greater than the vector of signals, $s_t$, at any $t$. This source of non-invertibility is what is typically assumed in the incomplete information literature. A different, and often more subtle reason, is when the elements in $\Gamma(z)$ combine in a way to create an “internal” source of non-invertibility. While the analysis of Section 2.2 formally illustrates this claim, the easiest way to see this is to consider the case of $m = n$, so that $\Gamma(z)$ is a square matrix. $\Gamma(z)$ is non-invertible for prediction purposes if its determinant vanishes at one or more points inside the unit circle. In this case, despite having the same number of shocks and signals, non-invertibility stems from the way signals combine over time and become themselves a source of noise.

We are interested in this internal source of non-invertibility, which we term “Confounding Dynamics”. Confounding dynamics naturally arise also when $m > n$, in which case they compound with the first source of non-invertibility, but their characterization can be substantially more involved. Because our focus is on showing that confounding dynamics can endogenously emerge in equilibrium, in the rest of the paper we focus mostly on the $m = n$ case. In Section 4.3 we analyze an example where the two sources of non-invertibility are simultaneously present, while we provide a more formal treatment of the $m > n$ case in Appendix B.3. The following definition formalizes the above discussion.

---

7Whittle (1983) shows that the computation of $\Gamma^*(L)$ is the dual to the solution of the $n^{th}$ order Ricatti equation in the Kalman filter approach to optimal linear-quadratic prediction. In both cases the objective is to figure out the variance-covariance of the optimal prediction errors. Hence, if the Ricatti equation is solved, then $\Gamma^*(z)$ can be determined; and if $\Gamma^*(L)$ is computed, then a solution to the Ricatti equation can be determined. In this paper we take the latter approach, while other authors, most notably Huo and Takayama (2016), take the former. Whether it is easier to solve the Ricatti equation or to obtain the canonical factorization often depends on the application at hand, and whether one is looking for an analytical or a numerical solution. The canonical factorization, however, has a wider scope of applicability. Whittle (1983) presents a generalized approach – which he terms the Hamiltonian approach – where the prediction problem is solved by the canonical factorization of a matrix function obtained by augmenting the plant and observation equations of a Kalman system with the saddle point conditions of an Hamiltonian optimization problem. Huertgen, Hoffmann, Rondina, and Walker (2016) make an explicit connection between these solution methodologies.
Definition CD. Let $s_t$ be specified as in (2.2), with $m = n$. The $s_t$ process is said to display confounding dynamics if there exists a $\lambda$, with $|\lambda| < 1$, such that $\det \Gamma(z)|_{z = \lambda} = 0$.

If the determinant vanishes at $z = |\lambda| \in (-1, 1)$, then it guarantees that $\Gamma(z)$ is not invertible and therefore knowledge of $s^t$ will not translate into direct knowledge of $u^t (u_t \neq \Gamma(L)^{-1}s_t)$. The Wold representation theorem implies that there always exists a $\Gamma^*(L)$ that is invertible. Deriving this $\Gamma^*(L)$ is the purpose of the next section.

The definition allows for an arbitrary number of roots at which the determinant of $\Gamma(L)$ vanishes. For simplicity, we will focus on characterizing equilibria with confounding dynamics due to only one root $\lambda$. An immediate consequence of this is that $\lambda$ must necessarily be real-valued. Appendix B.4 presents an example of confounding dynamics due to $N$ non-invertible roots. Taken together, the results in Appendix B.3 and B.4 ensure that the learning mechanism at the core of the equilibria that we characterize in Section 4 does not hinge upon the dimensions of $\Gamma(L)$ or the number of zeros inside the unit circle.

2.2 Prediction with Confounding Dynamics

To study the mechanism of confounding dynamics and showcase the usefulness of the Wiener-Kolmogorov prediction formula (2.4), we present a simple version of the prediction problem that operates at the heart of the rational expectations equilibria established in Section 4.

Let $Q(L) = 1$ in (2.1), so that $\omega_t = \varepsilon_t$, and $s_t$ is a univariate process specified as

$$s_t = -\lambda \varepsilon_t + \varepsilon_{t-1},$$

which results in $\Gamma(L) = (L - \lambda)\sigma_{\varepsilon}$, $u_t = \sigma_{\varepsilon}^{-1}\varepsilon_t$, and let $\lambda < 0$. Suppose that the prediction problem is to compute the mean-squared error minimizing prediction for $\varepsilon_t$ given that $s^t$ is observed. To fix ideas, imagine that $\varepsilon_t$ is the time-$t$ unobserved innovation in aggregate productivity in the economy, while $s_t$ is the observed market rental rate of physical capital. The prediction problem asks for an estimate of the current productivity innovation using the history of the market rental rate $s_t$.

To solve the problem we can apply the Wiener-Kolmogorov prediction formula (2.4) with $j = 0$. Note that $g_{\omega_s}(z) = \sigma_{\varepsilon}^2(z^{-1} - \lambda)$, whereas to obtain $\Gamma^*(L)$, we need to consider two possible cases. If $|\lambda| \geq 1$, $\Gamma^*(L) = \Gamma(L) = (L - \lambda)\sigma_{\varepsilon}$, which means that the stochastic process (2.7) is invertible, and therefore there exists a linear combination of current and past $s_t$’s that allows the exact recovery of $\varepsilon_t$. One can easily verify that applying (2.4) leads to $\Pi(L) = (L - \lambda)^{-1}$, and the optimal prediction corresponds to

$$P(\varepsilon_t|s^t) = \frac{s_t}{L - \lambda} = -\frac{1}{\lambda} \left(s_t + \lambda^{-1}s_{t-1} + \lambda^{-2}s_{t-2} + \lambda^{-3}s_{t-3} + \ldots \right) = \varepsilon_t,$$

which verifies that the history of $s^t$ contains all the information needed to perfectly know $\varepsilon_t$.

Consider now the case of $|\lambda| < 1$, so that $s_t$ displays confounding dynamics, according to Definition CD. Clearly, the prediction formula (2.8) is no longer well defined as the coefficients...
Rondina & Walker: Confounding Dynamics

diverge. The problem is that $\Gamma(L)$ is non-invertible at $L = \lambda$ and so our previous choice for $\Gamma^*(L)$ would not work in this case. In this simple environment, Rozanov (1967) shows that the canonical factorization is given by flipping the root outside of the unit circle, $\Gamma^*(L) = (1 - \lambda L)\sigma_\varepsilon$, which results in

$$
\Pi(L) = \sigma_\varepsilon^2 (L^{-1} - \lambda)(1 - \lambda L^{-1})^{-1} = -(1 - \lambda L)^{-1}.
$$

(2.9)

The optimal prediction for $\varepsilon_t$ is then

$$
P(\varepsilon_t|s_t) = -\lambda (s_t + \lambda s_{t-1} + \lambda^2 s_{t-2} + \lambda^3 s_{t-3} + ...) = -\lambda \left( \frac{L - \lambda}{1 - \lambda L} \right) \varepsilon_t,
$$

(2.10)

which results in a mean squared forecast error of $(1 - \lambda^2) \sigma_\varepsilon^2 > 0$, demonstrating that as $\lambda$ approaches one from below there is exact recovery of $\varepsilon_t$.

When the process is non-invertible, (2.10) shows that the history of current and past $s_t$’s reveals a particular linear combination of $\varepsilon_t$’s. Expanding this last term yields

$$
P(\varepsilon_t|s_t^t) = \lambda^2 \varepsilon_t - (1 - \lambda^2)[\lambda \varepsilon_{t-1} + \lambda^2 \varepsilon_{t-2} + \lambda^3 \varepsilon_{t-3} + \cdots].
$$

(2.11)

information + noise from confounding dynamics

Thus, the noise resulting from confounding dynamics takes an unusual form as it consists of a linear combination of past realizations of $\varepsilon_t$. This is the sense in which the non-invertibility of $\Gamma(L)$ is of an “internal” nature. Expression (2.11) suggests that the process (2.7) is informationally equivalent to a noisy signal about $\varepsilon_t$, where the noise is the linear combination of past shocks (in the bracketed term), and the signal-to-noise ratio is measured by $\lambda^2$. A $\lambda$ closer to zero results in less information and more noise, but, at the same time, it also makes past shocks less persistent. In fact as $\lambda \to 0$, there is no information in $s_t$ about $\varepsilon_t$ and the optimal prediction is 0, the unconditional average. As long as $|\lambda| \in (-1, 1)$, the value of $\varepsilon_t$ will never be learned and in this sense, the history of the fundamental shock acts as a standard noise shock. This is the defining characteristic of confounding dynamics.

To make the connection to the standard signal extraction problem more explicit, suppose that agents observe an infinite history of the signal

$$
x_t = \varepsilon_t + \eta_t,
$$

(2.12)

where $\eta_t \sim iid N(0, \sigma_\eta^2)$. The optimal prediction is well known and given by $P(\varepsilon_t|x^t) = \tau x_t$, where $\tau$ is the relative weight given to the signal, $\tau = \sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_\eta^2)$. It can be showed\(^9\) that the information content of (2.7) with $|\lambda| < 1$ is equivalent to (2.12), where equivalence is defined as equality of

---

\(^9\)See Online Appendix B.2 for a proof.
Figure 1: Impulse Responses of $x_t$ and $s_t$ to a one unit change in $\varepsilon_t$ for signal-to-noise ratios of $\tau = 1/2$, $\lambda = -1/\sqrt{2}$ (dashed, solid blue) and $\tau = 1/10$, $\lambda = -1/\sqrt{10}$ (dashed-dot, solid red). The black dotted line is the innovation itself, $\varepsilon_t$, normalized to 1.

variance of the forecast error conditioned on the infinite history of the observed signal, i.e.

$$
E \left[ (\varepsilon_t - P(\varepsilon_t|s_t))^2 \right] = E \left[ (\varepsilon_t - P(\varepsilon_t|x_t))^2 \right],
$$

when

$$
\lambda^2 = \tau.
$$

(2.13)

Notice that when the signal-to-noise ratio increases (decreases), this corresponds to a higher (lower) absolute value of $\lambda$. In the limit, as $\sigma_\eta^2 \to 0$, then $\lambda^2 \to 1$, which ensures exact recovery of the state in both cases.

While the informational content can be made identical, the dynamics of the two signal extraction problems are very different. To visualize this, we report the impulse response function for the prediction equations that contain confounding dynamics (2.11) and for the standard signal extraction problem (2.12) to a one time, one unit increase in $\varepsilon_t$ in Figure 1. We do this for both a low and high value of $\lambda^2$ (resp. $\tau$).

First notice that with confounding dynamics, (2.11) under-predicts the actual innovation on impact, with a smaller value of $\lambda$ under-predicting more significantly. This is due to the first term on the RHS of (2.11). The same is true for the standard signal extraction formulation (dashed lines). Agents weigh the initial innovation by the signal-to-noise ratio $\tau < 1$ and therefore under-predict on impact. This is where the similarities end. With confounding dynamics, periods two through six show waves of over- and under-prediction relative to the actual realization and relative to the standard signal extraction problem. As discussed above, the current and past innovations will persistently affect the prediction function several periods beyond impact. This
defining characteristic of confounding dynamics leads to the waves of over- and under-reaction. This is in contrast to the full information case and standard signal-extraction case where the impulse response is zero after impact. As already pointed out, the smaller the \( \lambda \), the larger the noise term in (2.11), but the less persistent the over- and under-prediction. Thus optimal signal extraction with confounding dynamics generates fluctuations where the full-information and exogenously imposed noise counterparts generate none. We now show how to embed these dynamics in a rational expectations model.

3 Model, Information, and Equilibrium

We now model confounding dynamics in a generic rational expectations formulation that permits many interpretations (e.g., monetary model, asset pricing model, etc.). We do this via dispersed information, which introduces well-known difficulties. We lay out a solution strategy that takes advantage of aforementioned mathematical properties and compare that strategy to alternative methodologies.

3.1 Model We consider models that are populated by a continuum of agents indexed by \( i \in [0, 1] \). Let \( \mu(i) \) be the density of agent \( i \) characterized by the information set at time \( t \), denoted by \( \Omega_{it} \). We are interested in the class of models in which the individual optimal choice can be represented by the dynamic expectational difference equation,

\[
\phi \mathbb{E} [X_{it+1} | \Omega_{it}] = \psi(L) X_{it},
\]

where

\[
X_{it} \equiv \left( x_{it} \ y_{t} \ \theta_{it} \right)^\top.
\]

Here \( \phi \equiv [\phi_x \ \phi_y \ \phi_{\theta}] \), is a vector of coefficients, and \( \psi(L) \equiv [\psi_x(L) \ \psi_y(L) \ \psi_{\theta}(L)] \), is a vector of square-summable lag polynomials in non-negative powers of \( L \). \( x_{it} \) is the choice variable under the control of the individual agent \( i \); \( y_t \) is an endogenous aggregate variable that agents take as given, and \( \theta_{it} \) is an exogenous stochastic process specified as the sum of an aggregate component \( \theta_t \) and an i.i.d. individual component \( v_{it} \). Formally

\[
\theta_{it} = \theta_t + v_{it}, \quad \text{where} \quad \theta_t = A(L) \varepsilon_t,
\]

with \( v_{it} \sim \mathcal{N}(0, \sigma_v) \), \( \varepsilon_t \sim \mathcal{N}(0, \sigma_{\varepsilon}) \), and \( A(L) \) is a square-summable polynomial in non-negative powers of \( L \). To close the model we need to specify a relationship between the distribution of \( x_{it} \) across agents, and the aggregate \( y_t \). We thus posit that

\[
\gamma(L) \int_0^1 X_{it} \mu(i) di = 0,
\]
where $\gamma(L) \equiv [\gamma_x(L) \, \gamma_y(L) \, \gamma_\theta(L)]$, is a vector of square-summable finite-degree lag polynomials in non-negative powers of $L$, and we assume $\gamma_x(L) \neq 0$.\(^{10}\) As we proceed with the analysis it will be useful to think of equation (3.1) as representing a demand (or supply) schedule for agent $i$, and (3.4) as the relative market clearing condition. However, the specific form depends on the particular application at hand. As we show in Section 3.3, this setup nests incomplete information models typical of the macro and finance literatures.

The expectational difference equation (3.1) is a dispersed information version of the system originally considered by Blanchard and Kahn (1980), and subsequently studied by Uhlig (1999), Klein (2000) and Sims (2002), among others. Dispersed information implies that individual expectations are heterogeneous, which implies that the aggregation in (3.4) will result in taking an average of expectations. In particular, model (3.1)-(3.4) can accommodate both average expectations of aggregate variables and average expectations of individual variables.

### 3.2 INFORMATION

In our dispersed information setup, we assume that the information set $\Omega_{it}$ of an arbitrary agent $i$ at time $t$ consists of the smallest closed subspace generated by the history of the random variable $\theta^t_i \equiv \{\theta_{it}, \theta_{it-1}, \ldots\}$, and the history of the aggregate variable $y^t = \{y_t, y_{t-1}, \ldots\}$. Specifically, $\Omega^t_i = \theta^t_i \vee y^t$, where the operator $\vee$ denotes the span (i.e., the smallest closed subspace which contains the subspaces) generated by the sequences $\theta^t_i$ and $y^t$. This notation simply suggests that expectations will be taken optimally; i.e., they will be consistent with the prediction formulas discussed in Section 2.2. In a multivariate moving-average setting, the invertible representation achieved via canonical factorization is the smallest closed subspace containing the observables, $\theta^t_i$ and $y^t$ (see Hoffman (1962)).

Given (3.1), $x_{it}$ will be a function of the history of idiosyncratic innovations, $v_{it}$, and the aggregate innovations, $\varepsilon_t$, namely

$$x_{it} = X(L)\varepsilon_t + V(L)v_{it}. \quad (3.5)$$

In addition, aggregation implies that $y_t$ is only a function of aggregate innovations, so that

$$y_t = Y(L)\varepsilon_t. \quad (3.6)$$

The signal structure can be thus represented as

$$\begin{pmatrix} \theta_{it} \\ y_t \end{pmatrix} = \Gamma(L) \begin{pmatrix} \sigma^{-1}_\varepsilon \varepsilon_t \\ \sigma^{-1}_v v_{it} \end{pmatrix}, \quad \Gamma(L) = \begin{bmatrix} A(L)\sigma_\varepsilon & \sigma_v \\ Y(L)\sigma_\varepsilon & 0 \end{bmatrix}. \quad (3.7)$$

We point out that our information set is in line with the typical information set assumed in the dispersed information rational expectations literature: we provide agents with both an exogenous signal about the aggregate unobserved state ($\theta_{it}$), and an endogenous signal that is determined

\(^{10}\)We make this assumption in order to keep the connection between (3.1) and (3.4) non-trivial. Allowing for $\gamma_x(L) = 0$, would imply that $y_t$ is directly determined by the process $\theta_t$, and, as a consequence, it would enter (3.1) as an exogenous variable, essentially duplicating the role of $\theta_t$ in that equation.
in equilibrium \((y_t)\). The analytical convenience of the signal structure \((3.7)\), for our purposes, is that the invertibility of the matrix \(\Gamma(L)\) hinges only upon the zeros of \(Y(L)\). At the same time, the structure imposes analytical discipline that is uncommon in the literature: the endogenous signal \(y_t\) can reveal perfectly the underlying state, under the appropriate parametrization of model \((3.1)-(3.4)\). Thus, the theorems below must establish both the degree to which information remains incomplete in equilibrium, along with the more standard existence and uniqueness conditions.

3.3 Examples

In this section we present four applications that can be cast into our model specification. The list is by no means exhaustive.

**Example 1:** Real Business Cycle with Capital. In a standard real business cycle model with capital, in presence of dispersed information about the aggregate productivity shock and incomplete insurance markets, the linear dynamics of capital around the steady state can be expressed as

\[
\alpha \beta E_{it} (k_{it+2}) + \eta (1 - \alpha \beta) E_{it} (r_{t+1}) - E_{it} (a_{it+1}) = \alpha (1 + \beta) k_{it+1} - \alpha k_{it} - a_{it},
\]

which is a standard second-order difference equation in capital, and where

\[
r_{t+1} = \int_0^1 a_{it+1} \mu(i) di - (1 - \alpha) \int_0^1 k_{it+1} \mu(i) di,
\]

is the market-clearing rental rate for capital. Here \(\beta\) is the subjective discount factor, \(\alpha\) is the capital share in the Cobb-Douglas output good technology, \(\eta\) is the elasticity of intertemporal substitution, and \(a_{it}\) is the individual productivity shock. \(^{11}\) Model \((3.8)-(3.9)\) maps into \((3.1)-(3.4)\) by setting \(x_{it} = k_{it+1}, y_t = r_t, \theta_{it} = a_{it},\) and

\[
\phi = \begin{pmatrix} \alpha \beta & \eta (1 - \alpha \beta) & -1 \end{pmatrix}, \psi(L) = \begin{pmatrix} \alpha (1 + \beta) - \alpha L & 0 & -1 \end{pmatrix}, \gamma(L) = \begin{pmatrix} (1 - \alpha) L & 1 & -1 \end{pmatrix}.
\]


\[
p^{*}_{it} = (1 - \beta \vartheta) (p_t + mc_{it}) + \beta \vartheta E_{it} (p^{*}_{it+1}),
\]

where \(p_t\) is the aggregate price level, defined as \(p_t = \vartheta p^{*}_{it} + (1 - \vartheta) p_{t-1}\), with \(p^{*}_{it} \equiv \int_0^1 p^{*}_{it} \mu(i) di\), and \(mc_{it}\) is the individual marginal cost at time \(t\) specified as \(mc_{it} = mc_t + v_{it}\) so that \(\int_0^1 mc_{it} \mu(i) di = mc_t\). The parameter \(\beta\) is the discount factor for price setters, while \(\vartheta\) measures the probability of resetting ones’ price in a given period. Define \(p_{it} \equiv \vartheta p^{*}_{it} + (1 - \vartheta) p_{t-1}\), which maintains \(\int_0^1 p_{it} \mu(i) di = \)

\(^{11}\)Equations \((3.8)\) and \((3.9)\) are derived as part of the application of Section 5. See that section for details.
$p_t$. The individual and aggregate price level dynamics can then be written as,

$$\beta \vartheta \mathbb{E}(p_{it+1} | \Omega_{it}) = p_{it} - \vartheta (1 - 2 \vartheta \beta + \beta) p_{it} + (1 - \vartheta)p_{it-1} - \vartheta (1 - \beta \vartheta)m_{ct}. \quad (3.11)$$

with

$$p_{it} = \int_0^1 p_{it} \mu(i) di. \quad (3.12)$$

Equations (3.11)-(3.12) map into (3.1)-(3.4) by setting $x_{it} = p_{it}$, $y_t = p_t$, $\theta_{it} = m_{ct}$, and

$$\phi = \begin{pmatrix} \beta \vartheta & 0 & 0 \end{pmatrix}, \psi(L) = \begin{pmatrix} 1 & -\vartheta (1 - 2 \vartheta \beta + \beta) + (1 - \vartheta)L & -\vartheta (1 - \beta \vartheta) \end{pmatrix}, \gamma(L) = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}.$$  

As recognized by Nimark (2008), in the presence of dispersed information a compact representation of the New Keynesian Phillips Curve cannot be obtained. However, once a solution for $p_t$ is derived from (3.11)-(3.12), inflation dynamics are immediately given by $\pi_t = p_t - p_{t-1}$.

**Example 3: Dynamic Asset Pricing.** Singleton (1987) presents a dynamic asset pricing model motivated by the market microstructure of the U.S. bond market, which features a competitive, Walrasian market structure with a single security that is traded among speculative investors and non-speculative or liquidity traders at the price $p_t$. The security is assumed to pay a constant coupon every period, which we normalize to zero. Purchases of the security are financed by borrowing at the constant rate $r$, and the wealth of investor $i$ evolves according to $w_{it+1} = z_{it} p_{it+1} - (1+r)(z_{it} p_{it} - w_{it})$. The $i^{th}$ investor is assumed to have a one-period investment horizon and to rank alternative investment strategies according to the utility $E_{it}\left[-\exp(-\varrho w_{it+1})\right]$, where $\varrho$ is the constant coefficient of absolute risk aversion. Singleton (1987) shows that the demand schedule for the risky asset takes the form

$$z_{it} = \frac{1}{\varrho \nu} \mathbb{E}_{it}(p_{t+1}) - \frac{1 + r}{\varrho \nu} p_t, \quad (3.13)$$

where $\nu$ is the variance of $p_{t+1}$ and is set to be an exogenous constant. Singleton (1987) assumes that the net supply of the asset, denoted by $n_t$, is specified as

$$n_t = f_t + \vartheta p_t. \quad (3.14)$$

The shock to net asset supply $f_t$ arises from nonspeculative traders (such as the U.S. Treasury, the Federal Reserve, financial intermediaries), that attempt to satisfy macroeconomic objectives for technical reasons related to the intermediation process. Nonspeculative traders are assumed to respond positively to an increase in prices; thus $\vartheta > 0$. Investors in setting their strategy $z_{it}$ are assumed to receive a private signal, $f_{it} = f_t + v_{it}$, about the shock to the net asset supply. Market clearing is therefore given by

$$\int_0^1 z_{it} \mu(i) di = n_t. \quad (3.15)$$

---

12 Several papers have since used a similar setup to study a broad range of asset pricing issues (e.g., Bacchetta and van Wincoop (2006)).
Model (3.13)-(3.15) maps into (3.1)-(3.4) by setting \( x_{it} = z_{it}, y_t = p_t, \theta_{it} = f_{it}, \) and
\[
\phi = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \psi(L) = \begin{pmatrix} \varphi & 1 + r & 0 \end{pmatrix}, \gamma(L) = \begin{pmatrix} 1 & -\vartheta & -1 \end{pmatrix}.
\]

**Example 4: Classical Monetary Models of Inflation.** In classical monetary models of inflation, money demand takes the form popularized by Cagan (1956),
\[
m_{it} - p_t = -\alpha \left( \mathbb{E}_{it} (p_{t+1}) - p_t \right), \tag{3.16}
\]
where \( m_{it} \) is nominal money demand by agent \( i \), \( p_t \) is an aggregate price index, and \( \alpha > 0 \). The money supply \( M_t \) is assumed to possess persistent dynamics specified as
\[
M_t = \rho M_{t-1} + f_t, \tag{3.17}
\]
where \( f_t \) is a money supply shock process. The money market clearing condition is then
\[
M_t = \int_0^1 m_{it} \mu(i)di \tag{3.18}
\]
Agents are assumed to receive a private signal, \( f_{it} = f_t + v_{it} \), about the money supply shock. Equations (3.16)-(3.18) map into (3.1)-(3.4) by setting \( x_{it} = m_{it}, y_t = p_t, \theta_{it} = f_{it} \), and
\[
\phi = \begin{pmatrix} 0 & -\alpha & 0 \end{pmatrix}, \psi(L) = \begin{pmatrix} 1 & -(1 + \alpha) & 0 \end{pmatrix}, \gamma(L) = \begin{pmatrix} 1 - \rho L & 0 & -1 \end{pmatrix}.
\]

### 3.4 Equilibrium Definition
Uncertainty is assumed to be driven by Gaussian innovations, which, together with linearity, implies that conditional expectations are computed as optimal linear projections. We thus have
\[
\mathbb{E}(X_{it+1}|\Omega_{it}) = \mathcal{P}[X_{it+1}|\Omega_{it}], \tag{3.19}
\]
and can apply the prediction formulas in Section 2.1 to compute conditional expectations. We are now ready to define a Rational Expectations Equilibrium for model (3.1)-(3.4).

**Definition REE.** A Rational Expectations Equilibrium (REE) is a stochastic process for \( \{X_{it}, i \in [0, 1]\} \) and a stochastic process for the information sets \( \{\Omega_{it}, i \in [0, 1]\} \) such that: (i) each agent \( i \), given her information set, forms expectations according to (3.19); (ii) \( \{X_{it}, i \in [0, 1]\} \) satisfies conditions (3.1)-(3.4).

The REE can be summarized by two statements: (a) given a distribution of information sets, there exists a market clearing distribution \( \{X_{it}, i \in [0, 1]\} \) determined by each agent \( i \)'s optimal prediction conditional on the information sets; (b) given a distribution \( \{X_{it}, i \in [0, 1]\} \), there exists a distribution of information sets that provides the basis for optimal prediction. Both statements (a) and (b) must be satisfied by the same distribution \( \{X_{it}, i \in [0, 1]\} \) and the same distribution of
information sets *simultaneously* in order to satisfy the requirements of a REE. This dual fixed point condition is standard in rational expectations with potentially heterogeneously informed agents and when endogenous variables convey information [see, Radner (1979) as an early example].

### 3.5 Weighted Sum of Expectations

Before discussing our solution methodology, we give a brief overview of the typical approach to solve model (3.1)-(3.4), which consists of two steps. The first step is to iteratively substitute the endogenous variables \(x_{it+j}\) and \(y_{it+j}\) forward by leading \((3.1)\) \(j\) periods forward and aggregating over agents. The end result is expressions for \(x_{it}\) and \(y_{it}\) that are a function of expectations of the exogenous variable \(\theta_t\) at all future horizons. The second step is then to compute those expectations, which is non-trivial due to the fact that the law of iterated expectations may not be operational. Most of the work that uses this approach rely on numerics to calculate these expectations.\(^{13}\)

Specifically, consider the univariate case, where \(y_t = \int_0^1 x_{it}\mu(i)di\), setting \(\phi_y = 0\), \(\psi_x(L) = 1\), \(\psi_y(L) = 0\), and \(\psi_y(L) = -1\). Next, bring (3.1) one period forward to obtain an expression for \(x_{it+1}\) and aggregate to get the analogue expression for \(y_{t+1}\). Substitute both expressions back into (3.1), which now will contain \(x_{it+2}\) and \(y_{it+2}\). Iterating on this procedure and aggregating over agents, the expression for \(y_t\) becomes\(^{14}\)

\[
y_t = \sum_{j=1}^{\infty} \mathbb{P}^j_{\phi_x} \left[ (\phi_x + \phi_y)^j \mathbb{E}^j(\theta_{t+j}) \right] \tag{3.20}
\]

Here \(\mathbb{E}^j(\theta_{t+j})\) stands for the \(j^{th}\) order average expectation of \(\theta_{t+j}\), with the convention that \(\mathbb{E}^0(\theta_t) = \theta_t\). Disregarding for a moment the operator \(\mathbb{P}^j_{\phi_x}\), equation (3.20) shows that \(y_t\) can be represented as a weighted sum of average expectations of higher order about the future realizations of the exogenous process \(\theta_t\). The higher the values of \(\phi_x\) and \(\phi_y\), the higher the relative weight of higher order thinking. The operator \(\mathbb{P}^j_{\phi_x}\) operates on the order of expectations \(j\) by reducing some of the higher order compounding depending on the position of \(\phi_x\) in the terms of the polynomial \((\phi_x + \phi_y)^j\). To see why the order might need to be reduced, consider the expression for \(\phi_x \mathbb{E}_{it}(x_{it+1})\) which contains the term \(\phi_x \mathbb{E}_{it+1}(x_{it+2})\), whose expression in turn contains \(\theta_{t+2}\). It follows that the law of iterated expectations (LIE) applies in this context so that \(\phi_x^2 \mathbb{E}_{it} \mathbb{E}_{it+1}(\theta_{t+2}) = \phi_x^2 \mathbb{E}_{it}(\theta_{t+2})\), and aggregation implies \(\phi_x^2 \mathbb{E}_{it}(\theta_{t+2})\) for \(j = 2\) in (3.20). Intuitively, in each round of the iterative substitutions to achieve representation (3.20) there are terms where agent \(i\) is taking expectations of both her own future expectations and of future average expectations. The law of iterated expectations applies to the former, so that the order of expectations is reduced, but not to the latter.\(^{15}\) It should be evident at this point that the second step required by the canonical approach—

---

\(^{13}\)Nimark (2010), and Melosi (2016) are recent examples of sophisticated numerical methods to characterize equilibria with dispersed information.

\(^{14}\)We provide the details of the derivation in Appendix B.5.

\(^{15}\)Mechanically, whether LIE applies or not at each iteration depends on the position of \(\phi_x\) in the coefficients of the polynomial \((\phi_x + \phi_y)^j\), i.e. on the set of permutations of size \(j\) of \(\phi_y\) and \(\phi_x\) with repetition. For instance, for the case of \(j = 2\), the set of terms that multiply \(\psi_0\) in (3.20) are \((\phi_y^2 + 2\phi_y\phi_x)\mathbb{E}_{t} \mathbb{E}_{t+1}(\theta_{t+2}) + (\phi_x\phi_y + \phi_x^2)\mathbb{E}_{t}(\theta_{t+2})\). For more details on this see Appendix B.5.
computing closed form solutions for the expectations of arbitrary order—is a daunting task under dispersed information. As already remarked and discussed thoroughly in the next section, we approach the solution to (3.20) from a completely different angle.

### 3.6 Solution Methodology

Our aim is to characterize an equilibrium for model (3.1)-(3.4) with confounding dynamics. From definition CD, the critical requirement for confounding dynamics to emerge is that the information matrix \( \Gamma(L) \) must be non-invertible at a \( \lambda \in (-1, 1) \). However, there is no guarantee that this condition will hold. Following Townsend (1983a), our approach is to formulate a guess for the endogenous variables that follows a generic polynomial in the underlying shocks, and then derive conditions on the exogenous parameters that yield non-invertibility in equilibrium. The following steps describe our procedure when looking for an equilibrium with confounding dynamics.

1. Specify the guesses for \( x_{it} \) and \( y_t \) as generic polynomials of underlying shocks

   \[
   x_{it} = X(L) \varepsilon_t + V(L)v_{it}, \quad \text{and} \quad y_t = Y(L) \varepsilon_t. \tag{3.21}
   \]

   where \( y_t \) has confounding dynamics, so that

   \[
   Y(\lambda) = 0, \quad \text{for} \quad \lambda \in (-1, 1). \tag{3.22}
   \]

2. Given the signal matrix \( \Gamma(L) \), obtain the canonical factorization form \( \Gamma^*(L) \) under (3.22).

3. Use \( \Gamma^*(L) \) in formula (2.4) together with the guesses in (3.21) to obtain the conditional expectations in (3.1).

4. Aggregate over agents according to (3.4) and use the relationship between \( X(L) \) and \( Y(L) \) to substitute \( X(L) \) with \( Y(L) \) in (3.1). Both the right hand side and the left hand side will now be lag polynomial operators in \( \varepsilon_t \) and \( v_{it} \), and will thus provide the fixed point conditions for \( Y(L) \) and \( V(L) \).

5. Derive conditions on exogenous parameters so as to ensure that the solution exists and is unique, and that there exists a \( |\lambda| < 1 \), verifying (3.22). Once \( Y(L) \) is solved for, use (3.4) to recover \( X(L) \).

Note that at no point in the solution procedure one needs to worry about higher-order expectations. The so-called “higher-order thinking” that complicates the iterative approach outlined in Section 3.5 is implicit in how the guess (3.21) combines with the information matrix \( \Gamma(L) \) to provide a closed form for the first order expectations in (3.1). As recognized by Townsend (1983a), by guessing a generic lag polynomial, the higher-order beliefs are built into the guess and we do not have to track these terms explicitly, although higher-order beliefs can be backed out of the solution in closed form. The same solution procedure is followed when we solve for an equilibrium with full information,

15
with the only difference that condition (3.22) is not imposed, and thus does not have to be verified, and the signal matrix $\Gamma(L)$ corresponds to full information so that prediction formula (2.5) applies.

4 EQUILIBRIUM WITH CONFOUNDING DYNAMICS

This section establishes the main result of the paper: the existence of a rational expectations equilibrium with confounding dynamics in a dispersed information environment.

4.1 FULL INFORMATION BENCHMARK

We consider first the case of Full Information to establish a useful benchmark and to show, in the simplest of settings, how our solution methodology works. We define Full Information as the case when every agent is endowed with perfect knowledge of the aggregate and her own idiosyncratic innovations history up to time $t$. Denoting the full information set by $\tilde{\Omega}_{it}$, the set is formally specified as

$$\tilde{\Omega}_{it} = v^t_i \lor \varepsilon^t.$$ (4.1)

Here, and in the following analysis, we assume that agents know that the equilibrium relationship is given by (3.1)-(3.4). When information is full, all agents share the same expectations about the future value of the exogenous process $\theta_t$. As a consequence, the entire structure of higher-order expectations in (3.20) collapses to the common first-order expectation so that

$$y_t = \sum_{j=0}^{\infty} (\phi_x + \phi_y)^j E_t(\theta_{t+j}).$$ (4.2)

Hansen and Sargent (1980) worked out a formula to express the discounted sum of future expectations—such as the one in (4.2)—in closed form, which since then has been known as the Hansen-Sargent formula. Here we show how to derive the Hansen-Sargent formula by applying the methodology of Whiteman (1983).\(^{16}\) Our main theorem extends the Hansen-Sargent formula to models with incomplete information.

Throughout the full information analysis we will maintain that $\phi_\theta = 0$, $\psi_\theta(L) = -1$, and consider market clearing (3.4) with $\gamma_\theta(L) = 0$, so that

$$\bar{\gamma}_y(L)y_t = \int_0^1 x_{it} \mu(i) di.$$ (4.3)

Here $\bar{\gamma}_y(L) \equiv \frac{\gamma_y(L)}{\gamma_x(L)}$, and we assume that $\gamma_x(0) \neq 0$, so that $\bar{\gamma}_y(0)$ is well defined.\(^{17}\)

We begin by guessing that the solution takes the form, $x_{it} = \mathcal{X}(L)\varepsilon_t + \mathcal{V}(L)v_{it}$, and $y_t = \mathcal{Y}(L)\varepsilon_t$, where $\mathcal{X}(L)$, $\mathcal{V}(L)$ and $\mathcal{Y}(L)$ are square-summable lag polynomial in non-negative powers

\(^{16}\)Whiteman (1983) provides a rigorous treatment of solving linear rational expectation models using the space of analytic functions. We rely on his theorems to establish existence and uniqueness of equilibria and provide an overview of his approach in Online Appendix C.

\(^{17}\)The solution under the general specification (3.1)-(3.4) is not instructive for our purposes here and has been relegated to Appendix A.1.
of \( L \). Under full information, direct application of the Wiener-Kolmogorov formula (2.5) provides expressions for the relevant expectational terms,

\[
\mathbb{E}_{lt}(x_{it+1}) = \left[ \mathcal{X}(L) - \mathcal{X}(0) \right] L^{-1} \varepsilon_t + \left[ \mathcal{V}(L) - \mathcal{V}(0) \right] L^{-1} v_{it},
\]

(4.4)

\[
\mathbb{E}_{lt}(y_{t+1}) = \left[ \mathcal{Y}(L) - \mathcal{Y}(0) \right] L^{-1} \varepsilon_t,
\]

(4.5)

\[
\mathbb{E}_{lt}(\theta_{t+1}) = \left[ A(L) - A(0) \right] L^{-1} \varepsilon_t.
\]

(4.6)

Substituting these expressions into equation (3.1), invoking \( \mathcal{X}(L) = \tilde{\gamma}_y(L) \mathcal{Y}(L) \), and aggregating over agents, we obtain an equation featuring \( \mathcal{Y}(L) \), while \( \mathcal{V}(L) \) is washed out by the aggregation process. After substitution, the \( \varepsilon_t \) terms can be dropped, both sides can be multiplied by \( L \), and the expression can be rearranged to solve for \( \mathcal{Y}(L) \) so that

\[
\mathcal{Y}(L) = \frac{\left( \phi_x \tilde{\gamma}_y(0) + \phi_y \mathcal{Y}(0) + A(L)L \right)}{\Phi(L)},
\]

(4.7)

where,

\[
\Phi(L) \equiv \phi_x \tilde{\gamma}_y(L) + \phi_y - (\psi_x(L)\tilde{\gamma}_y(L) + \psi_y(L))L.
\]

(4.8)

\( \Phi(L) \) is the characteristic polynomial that drives the auto-regressive behavior and the stationarity of \( y_t \). Expression (4.7) is not a solution because it features the endogenous constant \( \mathcal{Y}(0) \) on both sides, and, if we evaluate both sides at \( L = 0 \), the term \( \mathcal{Y}(0) \) drops out and cannot be solved for. As shown in Whiteman (1983), this indeterminacy is pinned down by the requirement that \( \mathcal{Y}(L) \) be square-summable (i.e. stationary), which is equivalent to the denominator polynomial of \( \Phi(L) \) not having any roots inside the unit circle. Thus, the constant \( \mathcal{Y}(0) \) must be set to remove roots inside the unit circle in \( \Phi(L) \).

Whiteman (1983)’s procedure requires three cases to be considered: [i.] if there are no roots of \( \Phi(L) \) inside the unit circle, then an infinite number of equilibria exist because \( \mathcal{Y}(0) \) cannot be set uniquely. [ii.] if there are multiple roots of \( \Phi(L) \) inside the unit circle, then no stationary equilibria exists. [iii.] and if there is exactly one root of \( \Phi(L) \) inside the unit circle, then \( \mathcal{Y}(0) \) can be set uniquely and a unique equilibrium exists. Thus we make the following assumption.

**Assumption \((S)\). The polynomial \( \Phi(L) \) has exactly one root inside the unit circle.**

It is important to note that Assumption \((S)\) is not a special case. It is the standard assumption necessary to yield a unique rational expectations equilibrium (e.g., Sims (2002)) and it immediately implies that \( \Phi(L) \) can be factorized as

\[
\Phi(L) = (\zeta - L)\bar{\Phi}(L),
\]

(4.9)

where \(|\zeta| < 1\), and \( \bar{\Phi}(L) \) has no roots inside the unit circle. Under assumption \((S)\), the constant \( \mathcal{Y}(0) \) can be chosen so to introduce a root in the numerator polynomial of (4.7) that cancels the
non-stationary root $\zeta$ at the denominator. To wit,

$$(\phi_x \tilde{y}(0) + \phi_y)Y(0) + A(\zeta)\zeta = 0. \tag{4.10}$$

Solving for $Y(0)$ and substituting the expression into (4.7), one finally obtains the solution

$$Y(L) = \frac{A(\zeta)\zeta - A(L)L}{\Phi(L)}, \tag{4.11}$$

where, by construction, the root $\zeta$ in the denominator is now canceled with a zero at $\zeta$ in the numerator.\(^{18}\)

Equation (4.11) is an instance of the Hansen-Sargent formula. To better understand the formula, again let $\tilde{y}(L) = 1$, $\psi_x(L) = 1$, $\psi_y(L) = 0$, so that $\zeta = \phi_x + \phi_y$, and note that $1/(\zeta - L) = -L^{-1}(1 + \zeta L^{-1} + \zeta^2 L^{-2} + ....)$. After some manipulation, one can write $y_t = Y(L)\varepsilon_t$ as

$$y_t = \sum_{j=0}^{\infty} \zeta^j \theta_{t+j} - A(\zeta) \sum_{j=1}^{\infty} \zeta^j \varepsilon_{t+j}. \tag{4.12}$$

Comparing this expression to (4.2) shows that the Hansen-Sargent formula turns the infinite sum of expectations about future $\theta_t$’s into the difference between the infinite sum of future $\theta_t$’s under perfect foresight (the first summation term), minus the innovations to those future realizations that are not known at time $t$ given the specified information set (the second summation term). In this sense, it is a true prediction formula. It takes the best guess if all information were available to the agents and subtracts off the precise linear combination of unknown elements that minimizes the agent’s forecast error.

We conclude this section by pointing out that, even though one does not need to solve for $V(L)$ to figure out the solution for $y_t$ (because all agents are equally informed), one can apply the same steps as above to obtain a closed form for $V(L)$. The agent-specific component $V(L)v_{it}$ determines the cross section distribution of $x_{it}$. The characteristic polynomial that drives the autoregressive behavior of $V(L)$ is $\phi_x(L) \equiv \phi_x - \psi_x(L)L$. In order to have the cross-sectional distribution well defined at any point in time, except possibly for the unit root limit, we assume the following.

**Assumption (s).** The polynomial $\phi_x(L)$ has exactly one root inside the unit circle.

We report the closed form solution to $V(L)$ in Appendix A.1.

### 4.2 Equilibrium with Confounding Dynamics: Main Theorem

In this section we state our main Theorem, which provides conditions under which a REE with Confounding Dynamics exists. As stated in Section 3.2, we specify the information set as

$$\Omega_{it} = x_{it}^l \lor y^l. \tag{4.13}$$

\(^{18}\)In this sense, (4.11) can be simplified further to cancel the two roots. However, unless one specifies a specific process for $A(L)$, and an explicit form for the polynomials in $\Phi(L)$, such cancellation cannot be undertaken algebraically.
Agents thus observe the entire history of the exogenous process $\theta_{it}$ up to time $t$, together with the history of the aggregate variable $y_t$. In addition, the signal structure (2.2) and the model equations (3.1)-(3.4) are both common knowledge across agents.

By Definition CD, in solving for $y_t = Y(L)\varepsilon_t$, we must find restrictions on exogenous parameters that ensure that there exists a $\lambda \in (-1, 1)$ such that $Y(\lambda) = 0$. The following theorem derives conditions for such a $\lambda$ to exist.

**Theorem 1.** Consider model (3.1)-(3.4) with Assumptions (S) and (s). Let the information sets be specified as in (4.13). There exists a Rational Expectations Equilibrium with Confounding Dynamics of the form, $y_t = Y(L)\varepsilon_t$, with

$$Y(L) = \mathcal{Y}(L) - (1 - \tau(\lambda))(1 - \lambda^2) \frac{A(\lambda)}{(1 - \lambda L)\Phi(L)}, \quad (4.14)$$

if there exists a $\lambda \in (-1, 1)$ that solves

$$\mathcal{Y}(\lambda)\Phi(\lambda) = (1 - \tau(\lambda))A(\lambda), \quad (4.15)$$

where $\mathcal{Y}(L)$ is the full information solution, $\tau(\lambda) = \frac{A(\lambda)\sigma^2}{A(\lambda^2)\sigma^2 + \sigma^2}$, $A(\lambda)$ is a function of $\lambda$ that depends only on exogenous parameters, and $Y(L)$ in (4.14) has no zeros inside the unit circle other than $\lambda$. If there exists only one $\lambda \in (-1, 1)$ solving (4.15), the equilibrium (4.14) is unique.

**Proof.** See Appendix A.2. \qed

Theorem 1 provides sufficient conditions for an equilibrium with confounding dynamics to exist. The existence condition is given by (4.15), while the functional form of the equilibrium is (4.14).

The form of (4.14) is intuitive when contrasted with the full information counterpart. As noted above, the standard Hansen-Sargent formula (4.12) subtracts off the particular linear combination of future values of $\varepsilon_t$ that minimize the agent’s forecast error. As described in Section 2.2, confounding dynamics implies that a particular linear combination of past values of $\varepsilon_t$ are never revealed to the agent. In order to make a direct comparison to the Hansen-Sargent formula (4.12), set $\gamma_y(L) = 1$, $\psi_x(L) = 1$, $\psi_y(L) = 0$, $\phi_\theta = 0$ and $\psi_\theta(L) = -1$. According to Theorem 1, the solution under confounding dynamics can be written as

$$y_t = \sum_{j=0}^{\infty} \zeta^j \theta_{t+j} - A(\zeta) \sum_{j=1}^{\infty} \zeta^j \varepsilon_{t+j} - (1 - \tau(\lambda))(1 - \lambda^2)A(\lambda) \sum_{j=0}^{\infty} \lambda^j \varepsilon_{t-j}. \quad (4.16)$$

The first two components on the right-hand side are exactly the Hansen-Sargent formula. The third component— represented by the weighted sum $\sum_{j=0}^{\infty} \lambda^j \varepsilon_{t-j}$—arises due to confounding dynamics and is similar to the prediction formula of Section 2.2. Agents do not observe the linear combination of shocks weighted by $\lambda$. Conditioning down implies that this linear combination will (optimally) be subtracted from the Hansen-Sargent full-information equilibrium. The extent to which the unknown past matters depends on the imprecision of the private signal $\theta_{it}$, measured by $1 - \tau(\lambda)$;
the imprecision stemming from confounding dynamics, measured by \( 1 - \lambda^2 \); and the fixed point constant \( \mathcal{A}(\lambda) \).

Equation (4.15) provides the condition for the existence of equilibrium (4.14). It is obtained by evaluating the right-hand side of (4.14) at \( \lambda \) and setting it equal to zero. By doing so, (4.15) is ensuring that once the conditioning down due to confounding dynamics is taken into account, the \( \lambda \) responsible for such conditioning down must indeed be a point in which the equilibrium function is non-invertible. Condition (4.15) takes an intuitive form from an informational point of view. Note first that the LHS, \( \mathcal{Y}(\lambda)\tilde{\Phi}(\lambda) \), corresponds to the moving average part of the full information solution (4.11), evaluated at \( \lambda \). Suppose for a moment that the RHS of (4.15) is set to zero. If a \( |\lambda| \in (0,1) \) satisfying the condition existed, it would mean that the equilibrium with confounding dynamics would take the same form as the full information equilibrium \( \mathcal{Y}(L) \). However, equation (4.16) shows that in presence of confounding dynamics the unknown past must be subtracted from the full information equilibrium, which would make the full information solution \( \mathcal{Y}(L) \) inconsistent with confounding dynamics. The implication of this observation is that whenever the RHS of (4.15) is made small enough, an equilibrium with confounding dynamics may fail to exist. In particular, as the noise-to-signal ratio in private information \( \sigma_v/\sigma_\varepsilon \) declines, the signal-to-noise ratio, \( \tau(\lambda) \), gets closer to one, and eventually leads to non-existence of an equilibrium with confounding dynamics. When restriction (4.15) is not satisfied, the solution is given by the full-information equilibrium (4.12).

We finally note that the auto-regressive factor in (4.14), \( 1/(1 - \lambda L) \), injects into the equilibrium dynamics of \( y_t \) the waves of over- and under-reaction depicted in Figure 1, which are the hallmark of signal extraction under confounding dynamics. In Section 5, in the context of a real business cycle model, we provide a description of how economic incentives can combine with the waves of over- and under-reaction to deliver the fixed-point condition (4.15).

4.3 Exogenous Noise

As discussed in Section 2.1, there are two ways to preserve heterogeneous information in equilibrium—by continually adding exogenous noise until the noise terms overwhelm all signals, and/or by proving that there exists a zero inside the unit circle of the equilibrium as is done in Theorem 1. These categories are not mutually exclusive. Combinations of the two can certainly exist. In this section we first show that the standard way of introducing exogenous aggregate noise will not lead to the characteristic over- and under-reaction of the impulse response which is the hallmark of confounding dynamics.\(^{19}\) We then show that confounding dynamics can coexist with superimposed exogenous noise, and when they do, the characteristic over- and under-reaction reemerges. For transparency, we work within the stylized version of the generic rational expectations model \([\tilde{\gamma}_y(L) = 1, \psi_x(L) = 1, \psi_y(L) = 0, \phi_x = 0, \phi_y = \zeta = \phi_y] \), and we modify

\(^{19}\)Superimposing exogenous noise is a common practice in most of the recent (and past) literature on dispersed information and aggregate fluctuations [e.g., Grossman and Stiglitz (1980), Wang (1993), Makarov and Rychkov (2012), Angeletos and La’O (2013)].
the private signal of an arbitrary agent \( i \) to,\(^{20}\)

\[
\varepsilon_{it} = \varepsilon_t + \nu_{it}. \tag{4.17}
\]

All agents also observe the endogenous variable with superimposed exogenous noise \( \tilde{\eta}_t \),

\[
\tilde{y}_t = y_t + \tilde{\eta}_t. \tag{4.18}
\]

The noise \( \tilde{\eta}_t \) is assumed to be of the form \( \tilde{\eta}_t = U(L)\eta_t \), where \( U(L) \) is a ratio of two lag polynomials in non-negative powers of \( L \), and \( \eta_t \) is i.i.d. Gaussian with distribution \( N(0, \sigma_\eta) \). Define the following relative signal-to-noise ratios,

\[
\tau_\eta \equiv \frac{1}{1 + \sigma_\varepsilon^2/\sigma_\eta^2 + \sigma_\varepsilon^2/\sigma_\eta^2}, \quad \tau_v \equiv \frac{1}{1 + \sigma_\varepsilon^2/\sigma_v^2 + \sigma_\varepsilon^2/\sigma_v^2}, \tag{4.19}
\]

and note that \( \lim_{\sigma_\eta \to \infty} \tau_v = \tau = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_\eta^2) \). Following our solution strategy, we posit a candidate solution \( y_t = Q_\varepsilon(L)\varepsilon_t + Q_\eta(L)\eta_t \). In order to achieve a closed-form solution, we follow Taub (1989) in specifying \( U(L) = Q_\varepsilon(L) - Q_\eta(L) \), such that \( \tilde{y}_t = Q_\varepsilon(L)(\varepsilon_t + \eta_t) \).\(^{21}\) The following proposition characterizes analytically a rational expectations equilibrium for the exogenous noise economy, without confounding dynamics.

**Proposition 1.** Consider model (3.1)-(3.4) with Assumptions (S) and (s) and let \( \bar{\gamma}_y(L) = 1 \), \( \psi_x(L) = 1 \), \( \psi_y(L) = 0 \), and \( \phi_x = 0 \), so that \( \zeta = \phi_y \). Let the information sets be specified as \( \Omega_{it} = \varepsilon_i^t \lor \tilde{y}^t \). Define \( \lambda(L) \equiv (LA(L) - \zeta \tau_v A(\zeta \tau_v))/(L - \zeta \tau_v) \), and let \( U(L) = \lambda(L) \). There exists a unique rational expectations equilibrium,

\[
y_t = \left( \lambda(L) + \zeta \tau_\eta \frac{\lambda(L) - \lambda(\zeta)}{L - \zeta} \right) \varepsilon_t + \zeta \tau_\eta \left( \frac{\lambda(L) - \lambda(\zeta)}{L - \zeta} \right) \eta_t. \tag{4.20}
\]

which does not yield confounding dynamics, provided that,

\[
\lambda(z) + \zeta \tau_\eta \frac{\lambda(z) - \lambda(\zeta)}{z - \zeta}, \tag{4.21}
\]

is invertible for all \( z \) inside the unit circle.

**Proof.** See Appendix A.3.

The form of the equilibrium (4.20) can be best understood by studying the limiting functions of the

\(^{20}\)Assuming the private signal is \( \varepsilon_{it} \), rather than \( \theta_{it} \), greatly simplifies the algebra in characterization of the equilibria of Propositions 1 and 2. All the key steps in the equilibrium derivation would go through if one were to consider \( \theta_{it} \).

\(^{21}\)The reasons for choosing a convenient form for \( U(L) \) are two-fold. First, it streamlines the analytical derivation of the canonical factorization of the variance-covariance matrix. Second, it ensures that a solution to the equilibrium exists that takes the form of a finite order ARMA representation, or, in the frequency domain jargon, of an analytic function that can be represented as the ratio of two polynomials. The conditions for the existence of an ARMA solution in presence of exogenous noise superimposed to endogenous variables is an open active area of research, see Huo and Takayama (2016).
noise terms. Note that the polynomial $\lambda(L)$ takes the form of a Hansen-Sargent formula involving $A(L)$ and $\tau_v \zeta$. To understand its role, suppose that the public information $\tilde{y}_t$ is made uninformative so that $\tau_\eta \to 0$ (i.e. $\sigma_\eta \to \infty$). The equilibrium would then just be equal to $y_t = \lambda(L)\varepsilon_t$, which is the first term in (4.20) with $\tau_v$ equal to $\tau$. As soon as public information is made informative two additional terms appear, one which captures the additional information about $\varepsilon_t$ transmitted by the public information, and the other that injects the public noise $\eta_t$ into the equilibrium price. Note that the two terms enter the equilibrium price with the same dynamics, which is a consequence of the assumption $U(L) = \lambda(L)$. This process is also characterized by a Hansen-Sargent formula involving $\lambda(L)$ and $\tau_\eta \zeta$. When public information is made arbitrarily precise, i.e. $\sigma_\eta \to 0$ so that $\tau_\eta = 1$ then (4.20) corresponds to the full information equilibrium (4.11).

A comparison with Theorem 1 reveals that the additional noise of Proposition 1 coming from (4.18) implies that condition (4.15), which guarantees heterogeneous beliefs are preserved in equilibrium, is no longer necessary. In fact, (4.21) is an explicit assumption that there are no zeros inside the unit circle, which is the standard assumption in models with exogenous noise. In turn, this implies that the equilibrium cannot support confounding dynamics.

Next we want to characterize a solution with both exogenous noise and confounding dynamics. To see this more clearly, suppose that $A(L) = 1 + \theta L$. Applying (4.11), the full information solution can be immediately obtained as the $MA(1)$ process,

$$y_t = (1 + \theta \zeta)\varepsilon_t + \theta \varepsilon_{t-1}. \quad (4.22)$$

Substituting into the equilibrium (4.20) under the assumption that, $\theta < 1/(1 - \zeta(\tau_v + \tau_\eta))$, so that the invertibility of (4.21) holds, yields,

$$y_t = (1 + \theta \zeta(\tau_v + \tau_\eta))\varepsilon_t + \theta \varepsilon_{t-1} + \zeta \tau_\eta (1 + \theta) \eta_t. \quad (4.23)$$

The impulse response dynamics of $y_t$ in (4.23) to a shock $\varepsilon_t$ are entirely consistent with the optimal prediction formula associated with the standard signal extraction problem described in Section 2.2. The impulse response to a shock in $\varepsilon_t$ is smaller than the full information counterpart at impact, since $\tau_v + \tau_\eta < 1$, but otherwise unchanged (i.e. it matches the dashed dynamics of Figure 1).

Next we want to characterize a solution with both exogenous noise and confounding dynamics. Under the same assumptions about the private and public information signals, we posit a candidate solution $y_t = \hat{Q}_\varepsilon(L)\varepsilon_t + \hat{Q}_{\eta}(L)\eta_t$, and let $U(L) = \hat{Q}_\varepsilon(L) - \hat{Q}_{\eta}(L)$. The following proposition holds.

**Proposition 2.** Consider model (3.1)-(3.4) and let $\tilde{\gamma}_y(L) = 1$, $\tilde{\gamma}_x(L) = 1$, $\tilde{\gamma}_\phi(L) = 0$, $\phi_x = 0$, so that $\zeta = \phi_y$. Let the information sets be specified as $\Omega_{it} = e_i^t \lor \tilde{y}_t^t$. Define $\tilde{\lambda}(L) \equiv (1 - \lambda L)(LA(L) - \lambda A(\lambda))/(L - \lambda)$, for some real constant $\lambda$. There exists a Rational Expectations Equilibrium with Confounding Dynamics of the form

$$y_t = \frac{L - \lambda}{1 - \lambda L} \left( \frac{(L - \zeta(1 - \tau_\eta))\tilde{\lambda}(L) - \tau_\eta \tilde{\lambda}(\zeta)}{(L - \zeta)(L - \zeta(1 - \tau_\eta)\tau)} \varepsilon_t + \tau_\eta \frac{(1 - \tau)\zeta \tilde{\lambda}(L) - (L - \zeta \tau_\eta)\tilde{\lambda}(\zeta)}{(1 - \tau)(L - \zeta)(L - \zeta(1 - \tau_\eta)\tau)} \eta_t \right), \quad (4.24)$$

$$= 0$$
if, and only if, there exists a \( \lambda \in (-1, 1) \) that solves

\[
(1 - \tau)\tilde{\lambda}((1 - \tau_n)\tau\zeta) + \tau\tau_n\tilde{\lambda} = 0.
\]

(4.25)

**Proof.** See Appendix A.4. □

The building block of this equilibrium is \( \tilde{\lambda}(L) \), which can be directly compared to the \( \lambda(L) \) function of Proposition 1. As discussed above, the second term of a Hansen-Sargent prediction formula has an informational interpretation in that it amounts to what must be subtracted away from a complete-information equilibrium. The conditioning down associated with \( \tilde{\lambda}(L) \) of Proposition 2 is due to the endogenous zero, \( \lambda \) determined by (4.25); while the conditioning down associated with \( \lambda(L) \) of Proposition 1 is due to the exogenous noise term, \( \tau_v \).

Most importantly, Proposition 2 reintroduces confounding dynamics. Under the specification

\[ A(L) = 1 + \theta L, \]

the equilibrium (4.24), when (4.25) is satisfied,\(^{22}\) is given by

\[
y_t = \frac{L - \lambda}{1 - \lambda L} \left( \theta - \lambda(1 + \lambda \theta) - \lambda \theta \zeta(\tau \eta + (1 - \tau \eta)) \right) \xi_{t-1} - \lambda \theta \xi_{t} - \zeta \tau \eta \lambda \theta \eta.
\]

(4.26)

Comparing expression (4.26) to (4.23), both contain an \( MA(1) \) term for \( \xi_t \), and a constant coefficient on \( \eta_t \). However, for equilibrium (4.26), the \( MA(1) \) term is multiplied by the factor \((L - \lambda)/(1 - \lambda L)\) which, as seen in Section 2.2, injects the dynamic pattern typical of confounding dynamics.\(^{23}\) The impulse response dynamics of \( y_t \) in (4.26) to a shock \( \xi_t \) is smaller than the full information counterpart at impact and it matches the qualitative behavior of confounding dynamics in Figure 1.\(^{24}\) Taken together, Propositions 1 and 2 show that it is the learning mechanism due to confounding dynamics, rather than the one due to exogenous noise, that injects persistence in innovations, and, simultaneously, an amplification pattern that resembles waves of optimism and pessimism.

### 5 Application: Business Cycle with Confounding Dynamics

In this section we apply our results to a model of business cycle fluctuations driven by productivity shocks. The purpose of this section is to analytically demonstrate the confounding dynamics mechanism within a well established framework. To achieve this goal, we work within a linearized model reminiscent of the islands model of Lucas (1975).

The economy consists of a continuum of islands indexed by \( i \in [0, 1] \). Each island is inhabited by an infinitely-lived representative household, and by a representative firm, also indexed by \( i \). Household \( i \) supplies labor services exclusively to firm \( i \) in a decentralized competitive labor market.

---

\(^{22}\)Equilibria (4.20) and (4.24) do not necessarily exist simultaneously for the same parameter values. However, our objective here is to compare the qualitative features of the equilibrium dynamics across the two different classes of equilibria they represent, one with exogenous noise only, the other with both exogenous noise and confounding dynamics. For such exercise, the space of existence across parameter values has a secondary relevance.

\(^{23}\)For \( \theta > 0 \), one can show that \( \lambda \in (-1, 0) \) when (4.25) is satisfied.

\(^{24}\)To see this one need to show that the impact coefficient in (4.26) is smaller than the impact coefficient in (4.22), which corresponds to \(-\lambda(\theta - \lambda(1 + \lambda \theta) - \lambda \theta \zeta(\tau \eta + (1 - \tau \eta)) \eta) < 1 + \theta \zeta\). This can be easily shown using the property that when \( \theta > 0 \), \(-\lambda \theta \in (0, 1)\).
or, equivalently, workers cannot move across islands. Households supply labor inelastically to firms, and the labor supply is normalized to 1. Households own capital in the economy, which is rented out to firms in a centralized spot market. Firms use capital and labor to produce output, also supplied in a centralized competitive spot market. Households derive utility from consuming the output good. Output is produced by firm $i$ according to a Cobb-Douglas technology with capital and labor inputs – with income shares $\alpha$, and $1 - \alpha$ respectively, and total factor of productivity that is firm-specific and denoted by $e^{a_{it}}$, where

$$a_{it} = a_t + v_{it}.$$  

The term $a_t$ is common across all the islands, while $v_{it}$ is a productivity component that is specific to island $i$. In what follows, we consider a log-linearized version of the model with full capital depreciation and constant elasticity of intertemporal substitution, denoted by $\eta > 0$.\(^{25}\) Household $i$ sets consumption intertemporally according to the Euler equation

$$\mathbb{E}_{it}(c_{it} - c_{it+1} + \eta r_{t+1}) = 0.$$  \(5.1\)

The intertemporal budget constraint is

$$(1 - \beta \alpha)c_{it} + \alpha \beta k_{it+1} = (1 - \alpha)w_{it} + \alpha r_t - \alpha k_{it},$$  \(5.2\)

where $k_{it+1}$ is the capital stock that household $i$ is carrying into period $t + 1$, $w_{it}$ is the wage rate, $r_t$ is the rental rate of capital, and $\beta \in (0, 1)$ is the subjective discount factor. The island-specific wage rate is given by, $w_{it} = \frac{1}{1-\alpha}(a_{it} - \alpha r_t)$. Aggregate capital is defined as $k_{t+1} \equiv \int_0^1 k_{it+1} \mu(i) di$, and market clearing implies an interest rate

$$r_t = a_t - (1 - \alpha)k_t.$$  \(5.3\)

Using the household’s budget constraint at $t$ and at $t+1$ to get expressions for $c_{it}$ and $c_{it+1}$, and leading (5.3) one period forward, one can substitute (5.1) into the Euler to obtain a second-order difference equation for capital $k_{it+1}$

$$\alpha \beta \mathbb{E}_{it}(k_{it+2}) + \eta (1 - \alpha \beta) \mathbb{E}_{it}(r_{t+1}) - \mathbb{E}_{it}(a_{it+1}) = \alpha (1 + \beta) k_{it+1} - \alpha k_{it} - a_{it},\quad (5.4)$$

which completely characterizes the equilibrium. As remarked in Section 3.3, the model maps into our general setting by specifying $x_{it} = k_{it+1}$, $y_t = r_t$, and $\theta_{it} = a_{it}$.

Finally, we assume that total factor productivity that is common across islands follows the $ARMA(1, 1)$ process

$$a_t = \rho a_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1},$$  \(5.5\)

so that $A(L) = \frac{1+\theta L}{1-\rho L}$, and with $\rho \in [0,1]$ and $\theta \geq 0$.

\(^{25}\)The fully specified model and the derivation of the log-linearization are reported in the Online Appendix B.6.
**Full Information** We first derive the full information ($\Omega_{it} = v_t \vee \varepsilon_t^j$) solution for aggregate capital and the interest rate. The full-information guess for island-specific capital is given by $k_{it+1} = K(L)\varepsilon_t + V(L)v_{it}$. From (5.3), the interest rate is immediately determined by $r_t = R(L)\varepsilon_t$ and where

$$R(L) = A(L) - (1 - \alpha K(L)L).$$

(5.6)

The characteristic polynomial associated with equation (5.4) can be determined as

$$\Phi(L) = \alpha \beta - (\eta(1 - \alpha \beta)(1 - \alpha) + (1 + \beta)\alpha) L + \alpha L^2 = \alpha(\zeta - L)(\beta/\zeta - L).$$

(5.7)

Given that $\alpha$ (capital’s share of production) and $\beta$ (subjective discount factor) are both less than one, (5.7) contains one root inside the unit circle ($\zeta$) and one outside ($\beta/\zeta$), and their product is always equal to $\beta$. That is, Assumption (S) holds with, $0 < \zeta < 1$, and, $\beta/\zeta > 1$.

Following the steps outlined in Section 4.1, the full information equilibrium for capital can be derived as the $AR(2)$ process

$$K(L) = \frac{\frac{\zeta}{\alpha \beta}(1 + \kappa + \theta L)}{(1 - \rho L)(1 - \frac{\zeta}{\beta} L)},$$

(5.8)

and the interest rate takes an $ARMA(2,2)$ form

$$R(L) = \frac{1 + \left(\frac{\zeta}{\alpha \beta}(1 + (1 - \alpha)\kappa)\right)L - \frac{\zeta}{\alpha \beta} \theta L^2}{(1 - \rho L)(1 - \frac{\zeta}{\beta} L)}.$$

(5.9)

where $\kappa \equiv \frac{(\theta + \rho)(1 - \zeta)(\alpha \beta / \zeta - 1)}{(1 - \rho \zeta)(1 - \alpha)}$. \(^{26}\)

Figure 2 reports the impulse responses of capital and the interest rate to a persistent unitary shock in aggregate productivity in the full information equilibrium (5.8)-(5.9). The parameters for the aggregate productivity process are set to $\rho \approx 1$ and $\theta = 1.5$. This calibration of the productivity process implies that a unitary shock to $\varepsilon_t$ raises $a_t$ by 1 on impact, by an additional 2.5 in the subsequent period, and no additional growth thereafter. More generally when $\theta > 1$, technical knowledge is assumed to diffuse slowly at first, then accelerate, only to level off in subsequent periods. This parameterization corresponds to a canonical “S-shaped” diffusion process. A productivity process with this property emerges when, for example, a new technology (e.g. a new computer chip) is first adopted only by the most advanced sectors in the economy, and then it diffuses to the rest of the economy (see Canova (2007), pages 115-116 for further discussion). We consider two values for elasticity of substitution, $\eta = 2$ and $\eta = 6$ in Figure 2. In both cases capital climbs smoothly towards the same new permanent level. However in the case of low elasticity ($\eta = 2$), the climb takes longer because with a low elasticity of substitution, agents are less willing to give up their current consumption to accumulate more capital, everything else equal.\(^{27}\)

The response of the interest rate exhibits the general equilibrium implication of a more sluggish...
Figure 2: Aggregate Productivity Shock with Full Information

Impulse response of Capital $k_{t+1}$, and Interest Rate $r_t$, to a unitary positive shock to aggregate productivity $a_t$ in the Full Information equilibrium. The dashed-black line represents the response for the equilibrium with $\eta = 2$; the solid-black line represents the response for the equilibrium with $\eta = 6$. Remaining parameter values are $\beta = 0.985$, $\alpha = 0.33$, $\rho \approx 1$, $\theta = 1.5$. The impulse responses for capital are normalized with respect to the impact response for the $\eta = 6$ case.

adjustment of capital. Under $\rho \approx 1$, one can show that, $R(L) \approx (1 + \frac{\zeta}{\alpha \beta} \theta L)/(1 - \frac{\zeta}{\beta} L)$. At impact the reaction of the interest rate is always 1: the capital stock is fixed, and the increase in the interest rate mimics the unitary shock to productivity. In the subsequent period, the response coefficient is $\frac{\zeta}{\beta} (1 + \frac{\alpha}{\beta})$, which, for $\eta$ low enough, can be bigger than 1. For $\eta = 6$, the root $\zeta$ is small, and the increase in the supply of capital is more than enough to reduce the interest rate after impact. In the $\eta = 2$ case, on the other hand, $\zeta$ is larger, and the small increase in the supply of capital keeps the interest rate from decreasing. The combined effect of a higher interest rate and a longer time to adjust finally brings down the interest rate in the subsequent periods. The contrasting forces at work in this example carry over to equilibria with confounding dynamics, to which we now turn.

**Confound Dyanmics** Because households participate in two competitive markets every period – the labor market and the rental market for capital – they observe the island-specific wage rate $w_{it}$, and the rental rate $r_t$. The observation of $w_{it}$ and $r_t$ implies that household $i$ can always back out $a_{it}$ at time $t$ through the expression for $w_{it}$ reported above. As a consequence, observing the prices of labor and capital is equivalent to the information set

$$\Omega_{it} = a_{it}^t \vee r^t. \quad (5.10)$$

We also assume that households cannot observe the aggregate capital $k_t$, so to avoid the full revelation of $a_t$, and thus $v_{it}$, which would be implied by (5.3).\(^{28}\) Following Theorem 1, existence of confounding dynamics requires that the process for $r_t = R(L)\varepsilon_t$, has the following property,

$$R(\lambda) = 0, \quad (5.11)$$

\(^{28}\)There are many other information structures that would preserve confounding dynamics in this setting and would be consistent with the general specification of Section 3.1.
for a \( \lambda \in (-1, 1) \). A direct application of Theorem 1 leads to the following corollary.

**Corollary 1.** Consider the Real Business Cycle model (5.3)-(5.5). Let the information sets be specified as in (5.10). There exists a Rational Expectations Equilibrium with Confounding Dynamics of the form, \( k_{t+1} = K(L)\varepsilon_t \), and \( r_t = R(L)\varepsilon_t \), with

\[
K(L) = K(L) - (1 - \tau(\lambda))C(\lambda) \frac{(1 - \frac{\zeta}{\theta}(1 - \lambda^2))}{(1 - \frac{\zeta}{\sigma}(1 - \lambda L))},
\]

and \( R(L) = A(L) - (1 - \alpha)K(L)L \), if there exists a \( \lambda \in (-1, 1) \), that solves

\[
\mathcal{R}(\lambda) = (1 - \tau(\lambda))(1 - \alpha)C(\lambda)\lambda,
\]

where \( C(\lambda) \equiv \frac{(1-\beta)((1-\lambda)(1-\lambda^2)\beta)+(1-\tau(\lambda))(1-\lambda^2)\beta(\frac{1-\alpha}{\beta}+\zeta)+(1-\lambda)\lambda(\frac{1+\alpha}{\beta}+\zeta)}{\lambda(1-\lambda)\beta(1-\lambda^2)\tau(\lambda)(\beta-\zeta)} \), \( K(L) \) and \( \mathcal{R}(L) \) are as in (5.8) and (5.9), \( \tau(\lambda) \equiv \frac{(1+\theta\lambda)^2\sigma_2^2+1-\rho\lambda^2\sigma_2^2}{(1+\theta\lambda)^2\sigma_2^2+1-\rho\lambda^2\sigma_2^2} \), and \( R(L) \) has no zeros inside the unit circle other than \( \lambda \). If there exists only one \( \lambda \in (-1,1) \) solving (5.13), the equilibrium (5.12) is unique.

While the functional forms of equations (5.12)–(5.13) have the same general structure as Theorem 1 (and same interpretation), the context of the application allows us to gain additional insights into the existence and behavior of an equilibrium with confounding dynamics. For the calibration discussed below (and for a large neighborhood around the calibration), \( \lambda \in (-1, 0) \) and \( C(\lambda) > 0 \). One immediate implication of \( C(\lambda) > 0 \) is that capital will under-react on impact compared to the full information case. The intuition of this result is consistent with our discussion in Section 2.2: under confounding dynamics, agents do not know exactly whether a positive innovation in aggregate productivity has been realized, and will act as if there is a chance that aggregate productivity is unchanged.

A second implication of \( C(\lambda) > 0 \) is that a necessary condition for an equilibrium with confounding dynamics to exist, i.e. (5.13) to be satisfied, is \( \mathcal{R}(\lambda) < 0 \). Table 1 reports the endogenous values of \( \lambda \), solved using (5.13), for several numerical combinations of the idiosyncratic noise in \( a_t \), \( \sigma_v/\sigma_\varepsilon \), the elasticity of substitution, \( \eta \), and the moving average parameter, \( \theta \). In Panel 1 we hold fixed \( \sigma_v/\sigma_\varepsilon = 1 \), and vary \( \eta \) and \( \theta \). Columns (a)-(c), show that, for \( \theta = 0.5 \), there is no equilibrium with confounding dynamics, independent of the value of the elasticity of substitution. Recall that with \( \rho \approx 1 \), we have \( \mathcal{R}(L) \approx \frac{1+\frac{\zeta}{\theta}L}{1-\frac{\zeta}{\theta}L} \), which results in the necessary condition

\[
\mathcal{R}(\lambda) < 0 \iff \lambda \leq -\frac{\alpha\beta}{\zeta}\theta < 0.
\]

Whenever \( \theta > 0 \) is too small, this condition cannot possibly hold for \( \lambda \in (-1, 0) \). This is consistent with our optimal prediction formulas, established in Section 2.2. Recall that if the moving average representation was fundamental with respect to the underlying innovations, then perfect revelation of the state was achieved. For values of \( |\theta| \in (0,1) \), the exogenous process \( (a_t) \) perfectly reveals
Table 1: Existence of Equilibrium with Confounding Dynamics

<table>
<thead>
<tr>
<th>Panel 1: ( \sigma_v/\sigma_\varepsilon = 1 )</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity of Substitution, ( \eta )</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Moving Average, ( \theta )</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Confounding Dynamics, ( \lambda )</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>-0.49</td>
<td>-0.61</td>
<td>-0.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel 2: ( \eta = 2 )</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise-Signal Ratio, ( \sigma_v/\sigma_\varepsilon )</td>
<td>0.1</td>
<td>0.5</td>
<td>2</td>
<td>0.1</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>Moving Average, ( \theta )</td>
<td>1.25</td>
<td>1.25</td>
<td>1.25</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Confounding Dynamics, ( \lambda )</td>
<td>None</td>
<td>-0.84</td>
<td>-0.83</td>
<td>-0.71</td>
<td>-0.56</td>
<td>-0.55</td>
</tr>
</tbody>
</table>

Existence of Equilibrium with Confounding Dynamics for Numerical Values of the noise-to-signal ratio in \( a_{it} \), \( \sigma_v/\sigma_\varepsilon \), the elasticity of intertemporal substitution, \( \eta \), and the moving average coefficient in \( a_t \), \( \theta \). The rest of the parameters are set at \( \beta = 0.985 \), \( \alpha = 0.33 \), \( \rho \approx 1 \). For the numerical combinations reported, when \( |\lambda| < 1 \) exists, it is also unique.

the underlying shocks, which gets embedded in endogenous variables. As \( \theta \) increases from 0.5 to 1.5 (columns (d)–(f)), the non-fundamental root in the exogenous process \( a_t \), coupled with the idiosyncratic uncertainty, is sufficient to prevent full revelation of the underlying shocks.

Panel 1 also shows that as the elasticity of substitution increases, the confounding dynamics root increases in absolute value. As agents become more sensitive to changes in the interest rate, the deviation in capital accumulation from the full information counterpart on impact diminishes. This can be seen directly in (5.14) through the endogenous root \( \zeta \), as it can easily be shown that \( \partial \zeta / \partial \eta < 0 \).

Panel 2 of Table 1 looks at the effect of changing the idiosyncratic noise in the private signal \( a_{it} \) on the existence of an equilibrium with confounding dynamics, for a fixed value of \( \eta = 2 \), and two values for \( \theta \): 1.25 and 2. Column (a) shows that if the information contained in the private signal is high \( \sigma_v/\sigma_\varepsilon = 0.1 \), an equilibrium with confounding dynamics does not exist, while it exists once the private signal is made less informative (columns (b) and (c)). This is consistent with our interpretation of condition (4.15) in Theorem 1, and it is also intuitive: with a very precise private signal agents react to aggregate productivity in a manner consistent with complete information, which makes the interest rate more informative, and thus prevents confounding dynamics. Finally, columns (b)-(c), and (d)-(f), show that the effect of increasing \( \sigma_v/\sigma_\varepsilon \) on confounding dynamics is highly non-linear, and tends to level-off quite rapidly: in response to a change from 0.5 to 2 to the noise-to-signal ratio, \( \lambda \) declines by merely 0.01 in absolute value.

We conclude the analysis by looking at the qualitative effects of confounding dynamics on capital and the interest rate. Figure 3 shows the response of capital, \( k_{t+1} \), and the interest rate, \( r_t \), to a persistent unitary positive shock to aggregate productivity \( a_t \) under confounding dynamics, measured in percentage deviation from the respective full information responses. We report the
responses for two values of the elasticity of substitution, $\eta = 2$ and $\eta = 6$, which result in $\lambda = -0.71$ and $\lambda = -0.85$, respectively. In both cases capital follows the typical over- and under-reaction pattern consistent with confounding dynamics. Capital under-reacts at impact (15\% and 10\% respectively), and then over-reacts after impact (4.5\% and 6\%), and then under-reacts again, and so on. For the $\eta = 2$ case the decay in fluctuations is faster, while for the $\eta = 6$ oscillations are more persistent. This is consistent with the impulses responses of prediction under confounding dynamics in Figure 1: for a lower absolute value of $\lambda$ the initial under-reaction is stronger, but the fluctuations are less persistent.

The key to sustaining confounding dynamics resides in the behavior of the interest rate, $r_t$, which is reported in the right panel of Figure 3. Because capital is predetermined, the interest rate under confounding dynamics increases at impact by 1, exactly as in the full information case. However, in the period after impact, due to the limited capital accumulation, it raises 20\% above full information. As agents cannot observe aggregate capital, the behavior of the interest rate “confuses” them in the following sense: in their learning effort, they input a realization for the interest rate in which a past innovation – the unitary jump in productivity – has a larger effect than under full information. They thus interpret the interest rate behavior as possibly implying a positive innovation in productivity in period $t = 1$. In other words, the innovation at impact becomes a source of noise in estimating the innovation in period 1. This translates into an overly-optimistic capital accumulation. In period $t = 2$, the resulting large increase in capital supply depresses the interest rate to a level (-10\% and -14\%) that is incompatible with the productivity innovation believed to exist in period 1. Beliefs about productivity turn then more pessimistic, which results in a slow-down in capital accumulation. The subsequent dynamics alternate the same mechanism at a gradually declining scale.

One interesting feature of the response in Figure 3 is that, despite the difference in the extent of
confounding dynamics across the two cases, the deviations from full information look remarkably similar in size for the first few periods. This reveals a fundamental tension that exists in the context of a real business cycle application between the severity of endogenous informational frictions and the impact they have on capital accumulation. From Table 1, we know that a lower elasticity of substitution results in stronger confounding dynamics, i.e. a less informative interest rate. However, a lower elasticity of substitution also implies that movements in the interest rate have smaller impact on capital accumulation. Indeed, it is for this exact reason that confounding dynamics are stronger. It follows that, even if the dynamics of the interest rate transmit little information on aggregate productivity, capital accumulation remains relatively unaffected because they are already sluggish in the first place. This is an important insight that extends beyond our application: in the presence of information transmitted by endogenous signals, the factors that hamper equilibrium information transmission can very well be the same factors that make information incompleteness less relevant for equilibrium allocations.

We believe that the numerical application of Figure 3 displays a qualitative behavior of capital that is interesting and promising for quantitative applications. First, we are able to generate dynamics that resemble waves of optimism and pessimism, just as a consequence of optimal learning from endogenous variables. Second, while the period of such waves is mechanically determined by the assumption that we look at equilibria with only one non-invertible root \( \lambda \), richer non-invertible conditions – such as ones with multiple roots, conjugate pairs, etc. – result in waves that can be longer and asymmetric in duration (we explore a simple example with multiple roots in Appendix B.4). Third, in order to keep things analytically tractable and transparent, we have assumed away additional sources of frictions, thereby limiting the potential of the model to provide quantitatively significant results. However, we envision a richer environment with several types of frictions, such as financial frictions – which are likely to introduce stronger sensitivity of allocations to the interest rate, or exogenous noisy signals, but where confounding dynamics remain a major determinant of equilibrium behavior.

Hur, Leeper, Rondina, and Walker (2016) apply Bayesian methods to estimate both a real business cycle model and a medium-sized new Keynesian model with and without confounding dynamics. Results suggest that allowing for confounding dynamics can substantially increase the empirical fit of these models relative to competing information specifications (e.g., news shocks, complete information).

6 Concluding Comments

We have introduced a rational expectations equilibrium that generates over- and under-reaction relative to its complete information counterpart. This systematic optimism and pessimism is generated from a simple and optimal learning mechanism that can be easily applied to any dynamic setting. Future work will seek to better understand the empirical properties of confounding dynamics by incorporating them into real and nominal business cycle models designed to be taken to data. Theoretical results of Section 5 and preliminary empirical results of Hur, Leeper, Rondina,
and Walker (2016) show much promise. Future work [Huertgen, Hoffmann, Rondina, and Walker (2016)] will also seek to show an equivalence between the analytic function approach advocated here and the more familiar time-domain approach. Contrasting these approaches in a side-by-side fashion will help to highlight the benefits of the analytic function approach while (hopefully) demystifying certain aspects of it.

References


A Proofs

A.1 Full Information Solution

This section derives the solution under full information for model (3.1)-(3.4). The fixed point condition under full information can be found by substituting (4.4)-(4.6) into (3.1), so that

\[
\phi_x [X(L) - Y(0)] L^{-1} \varepsilon_t + \phi_x [V(L) - V(0)] L^{-1} \varepsilon_t + \phi_y [Y(L) - Y(0)] L^{-1} \varepsilon_t + \phi_\theta [A(L) - A(0)] L^{-1} \varepsilon_t \\
= \psi_x (L) \phi_x (L) \varepsilon_t + \psi_x (L) V(L) \varepsilon_t + \psi_y (L) \psi_\theta (L) \varepsilon_t + \psi_y (L) A(L) \varepsilon_t + \psi_\theta (L) \varepsilon_t.
\]

(A.1)

This equation defines a fixed point condition for \( Y(L) \) with all the terms that multiply \( \varepsilon_t \). Collecting terms that multiply \( \varepsilon_t \), multiplying both sides by \( L \) and rearranging we get

\[
Y(L)(\phi_x + \psi_x (L)L) = \phi_x Y(0) + \psi_y (L)L.
\]

(A.2)

Note that \( \phi_x (L) \equiv \phi_x + \psi_x (L)L \), which, under assumption (s), has exactly one zero inside the unit circle, denoted by \( \zeta_x \). We thus pick \( \Psi(0) \) to remove such zero by setting

\[
\phi_x Y(0) + \psi_y (\zeta_x) \zeta_x = 0.
\]

(A.3)

Solving for \( \Psi(0) \), substituting back into (A.2) one finally obtains

\[
Y(L) = \frac{\psi_y (L)L - \psi_y (\zeta_x) \zeta_x}{\phi_x (L)}.
\]

(A.4)

We now focus on the fixed point for \( Y(L) \) and \( X(L) \). As remarked in the text, the fixed point condition does not feature any components of \( Y(L) \), so that one does not need to solve for the latter to obtain the former. To proceed with the solution there are two possibilities: solve for \( Y(L) \) and then recover \( X(L) \), or vice-versa. In general, both routes are possible, but there are situations in which one direction is substantially easier than the other. This depends on whether \( \gamma_x (0) \neq 0 \) or \( \gamma_y (0) \neq 0 \). We report here both cases. We first consider the case that works whenever \( \gamma_x (0) \neq 0 \). We begin by manipulating condition (3.4) to get the following relationship between \( X(L) \) and \( Y(L) \),

\[
X(L) = \tilde{\gamma}_y (L) Y(L) + \tilde{\gamma}_\theta (L) A(L),
\]

(A.5)

where \( \tilde{\gamma}_y (L) = - \frac{\gamma_y (L)}{\phi_x (L)} \), and \( \tilde{\gamma}_\theta (L) = - \frac{\gamma_\theta (L)}{\phi_x (L)} \). Using (A.5) to substitute for terms featuring \( X(L) \) in (A.1) one obtains

\[
\phi_x [\tilde{\gamma}_y (L) Y(L) - \tilde{\gamma}_y (0) Y(0)] L^{-1} \varepsilon_t + \phi_x [\tilde{\gamma}_\theta (L) A(L) - \tilde{\gamma}_\theta (0) A(0)] L^{-1} \varepsilon_t + \phi_x [V(L) - V(0)] L^{-1} \varepsilon_t \\
+ \phi_\theta [Y(L) - Y(0)] L^{-1} \varepsilon_t + \phi_\theta [A(L) - A(0)] L^{-1} \varepsilon_t + \phi_\theta [V(L) - V(0)] L^{-1} \varepsilon_t \\
= \psi_x (L) \tilde{\gamma}_y (L) Y(L) \varepsilon_t + \psi_x (L) \psi_\theta (L) A(L) \varepsilon_t \\
+ \psi_x (L) V(L) \varepsilon_t + \psi_y (L) \psi_\theta (L) A(L) \varepsilon_t + \psi_\theta (L) \varepsilon_t.
\]

(A.6)

Taking all the terms that multiply \( \varepsilon_t \) in (A.6), multiplying by \( L \) both sides and rearranging, one gets

\[
Y(L) \Phi(L) = Y(0)(\phi_x \tilde{\gamma}_y (0) + \phi_\theta) - \xi_y (L),
\]

(A.7)

where

\[
\xi_y (L) \equiv (\phi_x - \phi_\theta (L)L) \tilde{\gamma}_\theta (L) A(L) + (\phi_\theta - \phi_\theta (L)L) A(L) - (\phi_x \tilde{\gamma}_\theta (0) + \phi_\theta) A(0).
\]

(A.8)

Under assumption (s), \( \Phi(L) \) has exactly one zero inside the unit circle, denoted by \( \zeta \), which means that we can choose \( Y(0) \) to remove such zero. We thus set

\[
Y(0)(\phi_x \tilde{\gamma}_y (0) + \phi_\theta) - \xi_y (\zeta) = 0.
\]

(A.9)

Solving for \( Y(0) \), substituting into (A.7) and rearranging, one finally gets

\[
Y(L) = \frac{\xi_y (\zeta) - \xi_y (L)}{\Phi(L)}.
\]

(A.10)
The expression for $\mathcal{X}(L)$ can then be recovered using (A.5). Next we consider the case that works whenever $\gamma_y(0) \neq 0$. We begin by manipulating condition (3.4) to get the following relationship between $\mathcal{X}(L)$ and $\mathcal{Y}(L)$,

$$\mathcal{Y}(L) = \gamma_x(L)\mathcal{X}(L) + \gamma_0(L)A(L),$$

(A.11)

where $\gamma_x(L) = -\gamma_x(L)/\gamma_y(0)$ and $\gamma_0(L) = -\gamma_0(L)/\gamma_y(0)$. Using (A.11) to substitute for terms featuring $\mathcal{Y}(L)$ in (A.1) one obtains

$$\phi_x[\mathcal{X}(L) - \mathcal{X}(0)]L^{-1}\xi_t + \phi_x[V(L) - V(0)]L^{-1}v_{it} + \phi_y[\gamma_y(L)\mathcal{X}(L) - \gamma_y(L)\mathcal{X}(0)]L^{-1}\xi_t$$

$$+ \phi_y[\gamma_y(L)A(L) - \gamma_y(L)A(0)]L^{-1}\xi_t + \phi_y[A(L) - A(0)]L^{-1}\xi_t = \psi_x(L)\mathcal{X}(L)\xi_t + \psi_x(L)\mathcal{Y}(L)v_{it} + \psi_y(L)\gamma_x(L)\mathcal{X}(L)\xi_t$$

$$+ \psi_y(L)\gamma_y(L)A(L)\xi_t + \psi_y(L)\xi_t + \psi_y(L)A(L)v_{it} + \psi_y(L)\gamma_y(L)v_{it}.\tag{A.12}$$

Taking all the terms that multiply $\xi_t$ in (A.12), multiplying by $L$ both sides and rearranging, one gets

$$\mathcal{X}(L)\Phi_x(L) = \mathcal{X}(0)(\phi_x + \phi_y\gamma_y(0)) - \xi_x(L),\tag{A.13}$$

where

$$\Phi_x(L) = \phi_x + \phi_y\gamma_y(L) - \psi_x(L)L - \psi_y(L)\gamma_x(L)L,\tag{A.14}$$

and

$$\xi_x(L) \equiv (\phi_y - \psi_y(L)L)\gamma_y(L)A(L) + (\phi_y - \psi_y(L)L)A(L) - (\phi_y\gamma_y(0) + \phi_y)A(0).\tag{A.15}$$

Analogously to Assumption ($S$), let us assume that $\Phi_x(L)$ has exactly one zero inside the unit circle, denoted by $\hat{\zeta}$, which means that we can choose $\mathcal{X}(0)$ to remove such zero. We thus set

$$\mathcal{X}(0)(\phi_x + \phi_y\gamma_y(0)) - \xi_x(\hat{\zeta}) = 0.\tag{A.16}$$

Solving for $\mathcal{X}(0)$, substituting into (A.13) and rearranging, one finally gets

$$\mathcal{X}(L) = \frac{\xi_x(\hat{\zeta}) - \xi_x(L)}{\Phi_x(L)}\tag{A.17}.$$

The expression for $\mathcal{Y}(L)$ can then be recovered using (A.11).

### A.2 Proof of Theorem 1

**Step 1: Factorization** We operationalize the key requirement that $Y(\lambda) = 0$ for $\lambda \in (-1, 1)$ by specifying a guess of the form $Y(L) = (L - \lambda)G(L)$, where $G(L)$ has no zeros inside the unit circle. The first step in the proof is to then use the equilibrium guess to derive the canonical factorization for the information set, so that the Wiener-Kolmogorov formula (2.4) can be applied. The information set can be written as

$$
\begin{pmatrix}
\theta_{it} \\
y_{it}
\end{pmatrix}
= 
\begin{bmatrix}
A(L)\sigma_\epsilon & \sigma_v \\
(L - \lambda)G(L)\sigma_\epsilon & 0
\end{bmatrix}
\begin{pmatrix}
\tilde{\xi}_t \\
\tilde{v}_{it}
\end{pmatrix},
$$

(A.18)

where $\tilde{\xi}_t = \sigma_\epsilon\tilde{\xi}_t$, $v_{it} = \sigma_v\tilde{v}_{it}$, is a convenient normalization so that the variance-covariance matrix of the innovations vector is the identity matrix. It follows that

$$
\Gamma(L) = 
\begin{bmatrix}
A(L)\sigma_\epsilon & \sigma_v \\
(L - \lambda)G(L)\sigma_\epsilon & 0
\end{bmatrix}.
$$

(A.19)

The following Lemma shows the canonical factorization for $\Gamma(L)$.

**Lemma A1.** The canonical factorization $\Gamma^*(z)\Gamma^*(z^{-1})^T$ of the variance-covariance matrix $\Gamma(z)\Gamma(z^{-1})^T$, where $\Gamma(z)$ is defined in (A.19), is given by

$$
\Gamma^*(z) = \frac{1}{\sqrt{A(\lambda)\sigma_\epsilon^2 + \sigma_v^2}}
\begin{bmatrix}
A(z)A(\lambda)\sigma_\epsilon^2 + \sigma_v^2 & \sigma_v\sigma_\epsilon \frac{1 - \lambda z}{z} (A(z) - A(\lambda)) \\
A(\lambda)\sigma_\epsilon^2(z - \lambda)G(z) & \sigma_\epsilon G(z)(1 - \lambda z)
\end{bmatrix}.
$$

(A.20)
Proof. Using Rozanov (1967) procedure, \( \Gamma^*(z) \) is computed as

\[
\Gamma^*(z) = \Gamma(z) W_\lambda B_\lambda(z). \tag{A.21}
\]

where

\[
W_\lambda = \frac{1}{\sqrt{A(\lambda)^2 \sigma_v^2 + \sigma_v^2}} \begin{pmatrix} A(\lambda) \sigma_v & -\sigma_v \\ \sigma_v & A(\lambda) \sigma_v \end{pmatrix}, \quad \text{and} \quad B_\lambda(z) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1 - \lambda}{\lambda} \end{pmatrix}. \tag{A.22}
\]

The form of \( W_\lambda \) is obtained by application of Lemma C1 in the Online Technical Appendix. Solving out the matrix multiplication after some algebra one obtains (A.20).

**Step 2: Expectations** Equipped with the canonical factorization (A.20), we next derive the three expectation terms: \( \mathbb{E}_t(x_{it+1}) \), \( \mathbb{E}_t(y_{it+1}) \), and \( \mathbb{E}_t(\theta_{it+1}) \) (recall that \( \mathbb{E}_t(\theta_{it+1}) = \mathbb{E}_t(\theta_{it+1}) \)). The second and third in the list are given by

\[
\mathbb{E}_t \left( \begin{array}{c} \theta_{it+1} \\ y_{it+1} \end{array} \right) = \left[ L^{-1} \Gamma^*(L) \right]_+ \Gamma^*(L)^{-1} \left( \begin{array}{c} \theta_{it} \\ y_{it} \end{array} \right). \tag{A.23}
\]

Recalling that \( \left[ L^{-1} \Gamma^*(L) \right]_+ = \left[ \Gamma^*(L) - \Gamma^*(0) \right] L^{-1} \), and defining \( \tau(\lambda) = \frac{A(\lambda)^2 \sigma_v^2}{A(\lambda)^2 \sigma_v^2 + \sigma_v^2} \) one gets

\[
\mathbb{E}_t(\theta_{it+1}) = \left[ A(L) - A(0) \right] L^{-1} \varepsilon_t - (1 - \tau(\lambda)) \frac{\lambda^2}{\lambda(1 - \lambda)} \varepsilon_t - (1 - \tau(\lambda)) \frac{1}{\lambda} \left[ 1 - \frac{\lambda}{\lambda(1 - \lambda)} \right] v_{it}, \tag{A.24}
\]

\[
\mathbb{E}_t(y_{it+1}) = \left[ L^{-1} \Gamma^*(L) \right]_+ \Gamma^*(L)^{-1} \left( \begin{array}{c} \theta_{it} \\ y_{it} \end{array} \right). \tag{A.25}
\]

The term \( \mathbb{E}_t(\varepsilon_{it+1}) \) is substantially more involved to derive, due to the fact that the correlation between \( x_{it+1} \) and \( \theta_{it} \) exists not only because they both depend on \( \varepsilon_t \), but they also both depend on \( v_{it} \). Formally, the application of the Wiener-Kolmogorov formula leads to

\[
\mathbb{E}_t(x_{it+1}) = \left[ L^{-1} g_{x_i, \theta_i}(L) \left( \Gamma^*(L^{-1})^T \right)^{-1} \right]_+ \Gamma^*(L)^{-1} \left( \begin{array}{c} \theta_{it} \\ y_{it} \end{array} \right), \tag{A.26}
\]

where \( g_{x_i, \theta_i}(L) \) is the variance-covariance generating function between \( x_i \) and the information set. Given the equilibrium guess, such function takes the form

\[
g_{x_i, \theta_i}(L) = \begin{pmatrix} X(L)A(L^{-1})\sigma_v^2 + V(L)\sigma_v^2 & X(L)(L^{-1} - \lambda)G(L^{-1})\sigma_v^2 \\ 1 - \lambda(1 - \lambda) & \lambda(1 - \lambda) \end{pmatrix}. \tag{A.27}
\]

It follows that

\[
L^{-1} g_{x_i, \theta_i}(L) \left( \Gamma^*(L^{-1})^T \right)^{-1} = \left[ L^{-1} \left( V(L)\sigma_v^2 + X(L)\sigma_v^2 A(\lambda) \right) - \sigma_v \sigma_v L^{-1} \frac{1 - \lambda}{\lambda} \left( X(L) - V(L)A(\lambda) \right) \right]. \tag{A.28}
\]

The application of the annihilator operator requires to take the annihiland minus the principal part of its Laurent series expansion. All the terms have the usual principal part around \( L = 0 \). However, the term containing \( \frac{1 - \lambda}{\lambda} \) also has a principal part around \( L = \lambda \), it follows that

\[
\left[ \left( \frac{1 - \lambda}{\lambda} \right) \frac{\lambda}{L} (X(L) - V(L)A(\lambda)) \right]_+ = L^{-1} \left[ \left( \frac{1 - \lambda}{\lambda} \right) \left( X(L) - V(L)A(\lambda) \right) + \frac{\lambda}{\lambda(1 - \lambda)} (X(0) - V(0)A(\lambda)) \right] - \frac{1 - \lambda}{\lambda} \frac{\lambda^2}{L} (X(\lambda) - V(\lambda)A(\lambda)). \tag{A.29}
\]

Finally one gets

\[
\mathbb{E}_t(x_{it+1}) = L^{-1} \left[ X(L) - X(0) \right] \varepsilon_t - (1 - \tau(\lambda)) \frac{1 - \lambda}{\lambda(1 - \lambda)} \varepsilon_t + L^{-1} \left[ V(L) - V(0) \right] v_{it} + \sigma_v \sigma_v L^{-1} \frac{1 - \lambda}{\lambda(1 - \lambda)} \left[ X(\lambda) - V(\lambda) - (V(\lambda) - V(0))A(\lambda) \right] v_{it}. \tag{A.30}
\]

**Step 3: Fixed Point** We begin by manipulating condition (3.4) to get the following relationship between \( X(L) \) and
\[ Y(L), \quad X(L) = \tilde{\gamma}_y(L)Y(L) + \tilde{\gamma}_\theta(L)A(L), \tag{A.31} \]

where \( \tilde{\gamma}_y(L) = -\frac{\gamma_y(L)}{\gamma_y(L) + \gamma_\theta(L)} \), and \( \tilde{\gamma}_\theta(L) = -\frac{\gamma_\theta(L)}{\gamma_y(L) + \gamma_\theta(L)} \). Next we substitute the equilibrium guess and expressions (A.24), (A.25), and (A.30) into model (3.1), which leads to the expression

\[
\phi_x \left[ L^{-1} \left[ X(L) - X(0) \right] \varepsilon_t - (1 - \tau(\lambda)) \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ X(\lambda) - X(0) \right] \right] \varepsilon_t \\
+ L^{-1} \left[ V(L) - V(0) \right] \varepsilon_t + \frac{\tau(\lambda)}{\lambda(1 - \lambda)} \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ X(\lambda) - X(0) \right] \varepsilon_t \\
\phi_y \left[ [(L - \lambda)G(L) + \lambda G(0)]^{-1} \varepsilon_t - (1 - \tau(\lambda)) \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ A(L) - A(0) \right] \varepsilon_t - \tau(\lambda) \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ 1 - \frac{\lambda(0)}{\lambda(\lambda)} \right] \varepsilon_t \\
= \psi_x(L)(X(L)\varepsilon_t + V(L)\varepsilon_t) + \psi_y(L)(L - \lambda)G(L)\varepsilon + \psi_\theta(L)A(L)\varepsilon_t + \psi_\theta(L)\varepsilon_t. \tag{A.32} \]

As one would expect, both on the left and right hand sides there are lag polynomials that multiply \( \varepsilon_t \) and \( v_{it} \). Because the two stochastic process are uncorrelated, the equality must hold independently for the terms that multiply \( \varepsilon_t \) for those that multiply \( v_{it} \). Taking into account relationship (A.31), equation (A.32) thus defines two fixed points: one for \( (L - \lambda)G(L) \) and one for \( V(L) \). Differently from the full information case, the fixed point for the aggregate \( y_t \) (that defined by the terms multiplying \( \varepsilon_t \)) also contains elements of the function \( V(L) \), more precisely the constant \( V(0) - V(\lambda) \). Therefore, in order to solve for \( (L - \lambda)G(L) \), we need first to solve for \( V(L) \). Taking the fixed point condition for the terms that multiply \( v_{it} \), multiplying both sides by \( L \) and rearranging one obtains

\[
V(L)\phi_x(L) = \phi_xV(0) - \phi_x \frac{\tau(\lambda)}{\lambda(1 - \lambda)} \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ X(\lambda) - X(0) \right] \varepsilon_t \\
- \frac{\tau(\lambda)}{\lambda(1 - \lambda)} \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ \phi_yG(0) + \phi_\theta(A(\lambda) - A(0)) \right] L + \psi_\theta(L). \tag{A.33} \]

where \( \phi_x(L) \equiv \phi_x - \psi_x(L)L \). Similarly, the fixed point for \( (L - \lambda)G(L) \) is

\[
(L - \lambda)G(L)\Phi(L) = \\
- \phi_x \tilde{\gamma}_y(0)G(0) - \phi_x \tilde{\gamma}_\theta(A(L) - \tilde{\gamma}_\theta(0)A(0)) + \phi_x \left( 1 - \tau(\lambda) \right) \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ X(\lambda) - X(0) \right] \varepsilon_t \left( V(\lambda) - V(0) \right) A(\lambda) \]
\[
+ \phi_y \left[ \lambda - \left( 1 - \tau(\lambda) \right) \frac{1 - \lambda^2}{X(1 - \lambda)} \right] G(0) - \phi_\theta \left[ \left( A(L) - A(0) \right) - \left( 1 - \tau(\lambda) \right) \frac{1 - \lambda^2}{X(1 - \lambda)} \left[ A(\lambda) - A(0) \right] \right] L \\
+ \psi_x(L)\tilde{\gamma}_\theta(L)A(L)L + \psi_\theta(L)A(L)L. \tag{A.34} \]

where we have used (A.31) to substitute for, \( X(L) - X(0) \), and, \( X(L) \), and, \( \Phi(L) \equiv \phi_x(L) + \phi_y - \psi_y(L)L \). The next Lemma will prove very useful.

**Lemma A2.** \( V(\lambda) = \tilde{\gamma}_\theta(\lambda) \).

**Proof.** Evaluate (A.33) at \( \lambda \) and rearrange to obtain

\[
V(\lambda)\psi_x(\lambda) = \\
- \phi_x \frac{\tau(\lambda)}{\lambda(1 - \lambda)} \left[ X(\lambda) - X(0) \right] - \phi_x \left( 1 - \tau(\lambda) \right) \left( V(\lambda) - V(0) \right) \left( A(\lambda) - A(0) \right) + \psi_\theta(\lambda)\lambda. \tag{A.35} \]

Next, evaluate (A.34) at \( \lambda \) and rearrange to obtain

\[
0 = - \tau(\lambda)\phi_x \left( X(L) - X(0) \right) + \phi_x \left( 1 - \tau(\lambda) \right) \left( V(\lambda) - V(0) \right) A(\lambda) - \phi_y\tau(\lambda)G(0)\lambda - \phi_\theta \left( A(\lambda) - A(0) \right) \tau(\lambda) \]
\[
+ \psi_x(\lambda)\tilde{\gamma}_\theta(\lambda)A(\lambda)\lambda + \psi_\theta(\lambda)A(\lambda)\lambda. \tag{A.36} \]

Clearly, for (A.35) and (A.36) to hold, assuming \( A(\lambda) \neq 0, \psi_x(\lambda) \neq 0 \) and \( \lambda \neq 0 \), it must be that \( V(\lambda) = \tilde{\gamma}_\theta(\lambda) \). \( \square \)

We can now use Lemma A2 to substitute for \( V(\lambda) \) in (A.33) and (A.34). It follows that to solve for \( (L - \lambda)G(L) \) we
just need an expression for $V(0)$, to which we now turn. From assumption (s) we know that there is a root $\zeta_V$ that needs to be removed for $V(L)$ to be stationary. We achieve this by choosing the appropriate constant $V(0)$ so that the numerator on the right hand side of (A.33) vanishes when evaluated at $\zeta_V$.

$$\phi_xV(0) - \phi_x\frac{(-)^{(L)}}{A(\lambda)} \frac{1-\lambda^2}{(1-\lambda\zeta_V)} \left[ X(\lambda) - X(0) - (\gamma(\lambda) - V(0))A(\lambda) \right] \zeta_V$$

Using (A.31) so substitute for $\gamma(\lambda)$ so that $\gamma(\lambda) = \phi(\lambda) - A(\lambda) = A(0)$ and letting

$$m(\lambda) \equiv \frac{\tau(\lambda)(1-\lambda^2)\zeta_V}{(1-\lambda\zeta_V)\lambda - \tau(\lambda)(1-\lambda^2)\zeta_V},$$

and

$$n(\lambda) \equiv \frac{-\phi(\lambda)(1-\lambda^2)(A(\lambda) - A(0))\zeta_V - \phi_x\tau(\lambda)(1-\lambda^2)\gamma(0)A(0)\zeta_V - \psi(\zeta_V)\lambda A(\lambda)}{(1-\lambda\zeta_V)\lambda - \tau(\lambda)(1-\lambda^2)\zeta_V}.$$  

Next we used (A.38) in (A.34), and we also substitute $X(\lambda) - X(0)$ using (A.31) to get

$$(L - \lambda)G(L) = \frac{-\lambda G(0)(\phi_x\gamma(0) + \phi_y)H(L) + J(L)}{\Phi(L)(1-\lambda L)\lambda},$$  

where

$$H(L) = \lambda(1 - \lambda L) - (1 - \tau(\lambda))(1 - \lambda^2)(1 + m(\lambda))L,$$

and

$$J(L) = (1 - \tau(\lambda))(1 - \lambda^2)\left[ n(\lambda) - \phi_x\gamma(0)A(0) + \phi(\lambda) - A(0) \right] L + A(0)(\phi_x\gamma(0) + \phi_y)\lambda(1 - \lambda L)$$

$$- \left[ (\phi_x - \phi_y)(L)L\gamma(L) + \phi(\lambda) - \psi(\zeta_V)L \right] A(\lambda)\lambda(1 - \lambda L).$$

Under assumption (S), $\Phi(L)$ has a zero inside the unit circle at $\zeta$, which means that we need to choose the constant $G(0)$ so to cancel it. This is achieved by setting

$$-\lambda G(0)(\phi_x\gamma(0) + \phi_y)H(\zeta) + J(\zeta) = 0.$$  

Solving for $G(0)$ and substituting back into (A.41) one gets

$$(L - \lambda)G(L) = \frac{J(L)H(\zeta) - J(\zeta)H(L)}{\Phi(L)(1 - \lambda L)\lambda}.$$  

Next, recall that we defined

$$\xi_y(L) \equiv A(0)(\phi_x\gamma(0) + \phi_y) - \left[ (\phi_x - \psi(\zeta_V)L)\gamma(L) + \phi(\lambda) - \psi(\zeta_V)L \right] A(\lambda),$$

and letting

$$\hat{\xi} \equiv n(\lambda) - \phi_x\gamma(0)A(0) + \phi(\lambda) - A(0),$$

one can show that (A.45) can be written as

$$(L - \lambda)G(L) = \frac{\xi_y(L) - \xi_y(L)}{\Phi(L)} - (1 - \tau(\lambda))(1 - \lambda^2)(\zeta - L)\frac{\hat{\xi} - (1 + m(\lambda)\xi_y(\zeta)}{H(\zeta)\Phi(L)(1 - \lambda L)}.$$  

Using the factorization $\Phi(L) = (\zeta - L)\tilde{\Phi}(L)$, and defining

$$A(\lambda) \equiv \frac{\hat{\xi} - (1 + m(\lambda)\xi_y(\zeta)}{H(\zeta)},$$

39
expression (4.14) follows. Finally, for the solution to be consistent with the information set that we have used to derive it, it must be that the polynomial in (4.14) vanishes at $L = \lambda$, which corresponds to condition (4.15) in the Theorem.

The last step of the proof consists in making sure that when the equilibrium coefficients are evaluated using the $\lambda$ that solves (4.15), there are no other points at which $Y(L)$ vanishes inside the unit circle. More precisely, it has to be that there is no $\xi \neq \lambda$ that solves

$$\gamma(\xi)\tilde{\Phi}(\xi) = (1 - \tau(\lambda))(1 - \lambda^2)\frac{\mathcal{A}(\lambda)}{1 - \lambda^2},$$

(4.50)

such that $|\xi| \in (-1, 1)$. If this was not the case, then the information conveyed by $y_t$ in equilibrium would be inconsistent with the information used to derive the expectations that we use to determine the fixed point. More precisely, the factorization of $\Gamma(L)$ would be incorrect, as $\Gamma^*(L)$ in (A.20) would still be non-invertible. To see this, suppose that $\lambda$ is a solution to (4.15), while $\xi$ is a solution to (4.50), and they are both inside the unit circle. Then, the equilibrium function must have the form $G(L)(L - \lambda)(L - \xi)$, but the factorization above only removes the zero associated with $\lambda$. It follows that

$$\Gamma^*(L) = \frac{1}{\sqrt{A(\lambda)^2 + \sigma^2}}\begin{pmatrix} A(L)A(\lambda)\sigma^2 + \sigma^2 & \sigma\sigma_1\frac{A(L) - A(\lambda)}{(L - \xi)(L - \lambda)} \\ A(L)\sigma^2(\lambda - L)(L - \xi)G(L) - \sigma\sigma_2G(L) & \sigma\sigma_1G(L)(L - \xi)(1 - \lambda L) \end{pmatrix},$$

(4.51)

whose determinant still vanishes at $L = \xi$, so that $\Gamma^*(L)$ is not the appropriate factorization. In this case one can modify the initial guess and consider $N > 1$ roots inside the unit circle, looking then for a condition like (4.15) to deliver exactly $N$ solutions. We restrict our attention to $N = 1$ for simplicity and because the full description of the space of REE with confounding dynamics is beyond the scope of this paper, but we hope it is clear that our methods extend to the more general case.

**Step 4: No Information from the Model** The last thing to check to complete the proof is to ensure that there is no information that is transmitted by a clever manipulation of the model conditions – which are part of the information set of the agents – combined with the knowledge of the history of $\theta_{it}$ and $y_t$. For instance, suppose that the market clearing condition (3.4) is specified so that $\int_0^1 x_{it}(i)di = y_t$, which means that $y_t$ is the aggregate of $x_{it}$, then this would imply $X(L) = Y(L)$, which would result in $x_{it} - y_t = V(L)v_{it}$. Because rational agents know all this, they know that the difference $x_{it} - y_t$ is just a linear combination of the individual innovations $v_{it}$. It follows that they could, in principle, back out the realizations of $v_{it}$’s by inverting $V(L)$. More generally, the link between $X(L)$ and $Y(L)$ due to (3.4) can be used by rational agents to obtain additional information on the underlying innovations. For this not to happen, if one augments the information set of the agents by $x_{it} - y_t$, the information matrix must still be non-invertible at $\lambda$. The following Lemma shows that this is indeed the case for the equilibrium of Theorem 1.

**Lemma A3.** In the equilibrium with confounding dynamics of Theorem 1, consider the augmented information matrix $\tilde{\Gamma}(L)$, where

$$\begin{pmatrix} \theta_{it} \\ y_t \\ x_{it} - y_t \end{pmatrix} = \tilde{\Gamma}(L)\begin{pmatrix} \varepsilon_t \\ v_{it} \end{pmatrix} = \begin{pmatrix} A(L) & 1 \\ Y(L) & 0 \\ X(L) - Y(L) & V(L) \end{pmatrix}\begin{pmatrix} \varepsilon_t \\ v_{it} \end{pmatrix}.$$

(A.52)

The 2-by-2 minors of $\tilde{\Gamma}(L)$ all vanish at $\lambda$.

**Proof.** Matrix $\tilde{\Gamma}(L)$ has three minors, whose determinants are, respectively, $Y(L)$, $Y(L)V(L)$, and, $A(L)V(L) - (X(L) - Y(L))$. The first two minors clearly vanish at $\lambda$ since, by construction, $Y(\lambda) = 0$. For the third minor, use (A.31) to write

$$A(L)V(L) - (X(L) - Y(L)) = A(L)V(L) - \tilde{\gamma}_y(L)Y(L) - \tilde{\gamma}_{\theta}(L)A(L) + Y(L).$$

(A.53)

We thus need to show that

$$A(L)V(\lambda) = \tilde{\gamma}_{\theta}(\lambda)A(\lambda),$$

(A.54)
but this follows immediately from Lemma A2.

A.3 Proof of Proposition 1  

The first step in the proof is to recognize that the expectations of agent $i$ in equilibrium take the form of a linear combination of current and past realizations of the observed variables $\varepsilon_t$ and $\hat{\eta}_t$,

$$E_i(y_{t+1}) = \hat{\pi}_1(L)\varepsilon_t + \hat{\pi}_2(L)\hat{\eta}_t,$$

(A.55)

where $\hat{\pi}_1(L)$ and $\hat{\pi}_2(L)$ are assumed to be representable as ratios of two finite-degree lag polynomials with zeros outside the unit circle. The guess for the equilibrium price is specified as

$$y_t = Q_e(L)\varepsilon_1 + Q_o(L)\eta_t.$$

(A.56)

Substituting (A.55) in the equilibrium equation (3.1) and rearranging one obtains

$$y_t = \frac{\pi_1(L)}{1 - \pi_2(L)}\varepsilon_t + \frac{\pi_2(L)}{1 - \pi_2(L)}\eta_t,$$

(A.57)

where $\pi_1(L) \equiv \beta\hat{\pi}_1(L) + A(L)$, and $\pi_2(L) \equiv \beta\hat{\pi}_2(L)$. Recalling that $\hat{y}_t = y_t + \hat{\eta}_t$, the information set of agent $i$ can be then expressed as

$$\left(\begin{array}{c}
\varepsilon_t \\
\hat{y}_t
\end{array}\right) = \left(\begin{array}{cc}
1 & 1 - \pi_2(L) \\
1 - \pi_2(L) & 0
\end{array}\right) \left(\begin{array}{c}
\varepsilon_t \\
v_{it}
\end{array}\right) = \Xi(L) \left(\begin{array}{c}
\varepsilon_t \\
v_{it}
\end{array}\right).$$

(A.58)

Denoting the entire history of the signal vector by $\omega$, under the assumption that $U(L) = \pi_1(L)$ the variance-covariance generating function for the signal vector is

$$g_{ss}(z) = \left(\begin{array}{cc}
s_{\varepsilon}^2 + s_{\eta}^2 & \frac{\pi_1(z^{-1}) - \pi_2(z^{-1})}{1 - \pi_2(z)}
\\
\frac{\pi_1(z)}{1 - \pi_2(z)} & s_{\eta}^2
\end{array}\right).$$

(A.59)

In addition, the covariance generating function between the signal vector and $y$, the variable to be predicted, can be written as

$$g_{ys}(z) = \left(\begin{array}{cc}
\frac{\pi_1(z)}{1 - \pi_2(z)} & s_{\eta}^2
\\
\frac{\pi_1(z)\pi_2(z)}{1 - \pi_2(z)} & s_{\eta}^2 + s_{\varepsilon}^2 + s_{\eta}^2
\end{array}\right).$$

(A.60)

Applying the usual Wiener-Kolmogorov prediction formula one sees that

$$\left[\begin{array}{c}
\hat{\pi}_1(L) \\
\hat{\pi}_2(L)
\end{array}\right] = \left(\begin{array}{c}
L^{-1}g_{ys}(L)\Xi^*(L^{-1})^{-1}
\\
\Xi^*(L^{-1})^{-1}
\end{array}\right) + \Xi^*(L)^{-1},$$

(A.61)

where $\Xi^*(z)$ is the canonical factorization of the variance-covariance matrix $g_{ss}(z)$ such that $g_{ss}(z) = \Xi^*(z)\Xi^*(z^{-1})^T$.

Following Rozanov (1967) and Taub (1989), the factorization can be shown to take the form

$$\Xi^*(z) = \left(\begin{array}{cc}
s_{\varepsilon} & \frac{s_{\eta}^2}{s_{\varepsilon}^2 - s_{\eta}^2}
\\
0 & \frac{s_{\eta}^2}{s_{\varepsilon}^2 - s_{\eta}^2}
\end{array}\right),$$

(A.62)

where $s_{\eta}^2 \equiv s_{\varepsilon}^2 + s_{\eta}^2$, and $s_{\varepsilon}^2 \equiv s_{\varepsilon}^2 + \frac{s_{\eta}^2 s_{\varepsilon}^2}{s_{\varepsilon}^2 + s_{\eta}^2}$. Using (A.60) and (A.62) one can show that

$$\hat{\pi}_1(L) = \frac{s_{\eta}^2 s_{\varepsilon}^2}{s_{\eta}^2} \left(\begin{array}{c}
\pi_1(L) \\
\pi_2(L)
\end{array}\right) \frac{1 - \pi_2(L)}{L},$$

(A.63)

and

$$\hat{\pi}_2(L) = \frac{1 - \pi_2(L)}{\pi_1(L)} \left[\frac{s_{\eta}^2}{s_{\varepsilon}^2} \left(\begin{array}{c}
\pi_1(L) \\
\pi_2(L)
\end{array}\right) \frac{1 - \pi_2(L)}{L} \left(\begin{array}{c}
\pi_1(0) \\
\pi_2(0)
\end{array}\right) \left(\begin{array}{c}
\pi_1(0) \\
\pi_2(0)
\end{array}\right) \frac{1 - \pi_2(0)}{L(1 - \pi_2(0))} \right] + \frac{1}{L}. $$

(A.64)

Given our definition of $\pi_1(L)$, equation (A.63) is a fixed point equation that can be solved independently for $\hat{\pi}_1(L)$. 


Define $\tau_\nu \equiv \frac{\sigma^2_{\nu} \sigma^2_{\tau}}{\sigma^2_{\nu} \sigma^2_{\tau} + \sigma^2_{\tau} \sigma^2_{\nu} + \sigma^2_{\nu} \sigma^2_{\nu}} = \frac{\sigma^2_{\nu} \sigma^2_{\tau}}{\sigma^2_{\nu} \sigma^2_{\tau} + \sigma^2_{\tau} \sigma^2_{\nu} + \sigma^2_{\nu} \sigma^2_{\nu}}$, the fixed point condition results in

$$\hat{\pi}_1(L) = \frac{1 - \beta \hat{\pi}_2(L) + A(L) - A(0)}{L - \tau_\nu \beta}. \quad (A.65)$$

Because $\tau_\nu \beta < 1$, to ensure covariance-stability of $\hat{\pi}_1(L)$ we need to pick $\hat{\pi}_1(0)$ so to cancel the unstable root at the denominator. This is achieved by setting $\hat{\pi}_1(0) = \beta^{-1}(A(\tau_\nu \beta) - A(0))$. Substituting this into (A.65), a closed form solution for $\hat{\pi}_1(L)$ is obtained. Using $\pi_1(L) = \beta \hat{\pi}_1(L) + A(L)$ one finally obtains

$$\pi_1(L) = \frac{LA(L) - \tau_\nu \beta A(\tau_\nu \beta)}{L - \tau_\nu \beta}. \quad (A.66)$$

We let $\lambda(L) \equiv \pi_1(L)$ which agrees with the statement of Proposition 1. Turn now to condition (A.64). Using $\pi_2(L) = \beta \hat{\pi}_2(L)$ the fixed point condition can be expressed as

$$\frac{1}{1 - \beta \pi_2(L)} = \frac{1}{\pi_1(L)} \left( \frac{L \pi_1(L) - \beta \kappa(L)}{L - \beta} \right). \quad (A.67)$$

where

$$\kappa(L) \equiv \frac{\sigma^2_{\pi}}{\sigma^2_{\nu}} (\pi_1(L) - \pi_1(0)) + \frac{\pi_1(0)}{1 - \beta \pi_2(0)}. \quad (A.68)$$

Note that $\kappa(L)$ is a known function except for the constant $\hat{\pi}_2(0)$. Since $\beta < 1$, in order for the left hand side of (A.67) to be covariance-stationary, the right hand side should vanish at $L = \beta$. The constant $\hat{\pi}_2(0)$ can be conveniently chosen to achieve this by setting $\pi_1(\beta) - \kappa(\beta) = 0$. Solving this condition for $\hat{\pi}_2(0)$ and plugging the expression back into (A.67) one obtains

$$\frac{1}{1 - \beta \pi_2(L)} = 1 + \beta \tau_2 \frac{\pi_1(L) - \pi_1(\beta)}{\pi_1(L)(L - \beta)} \quad (A.69)$$

where $\tau_2 \equiv \frac{\sigma^2_{\nu} \sigma^2_{\tau}}{\sigma^2_{\nu} \sigma^2_{\tau} + \sigma^2_{\tau} \sigma^2_{\nu} + \sigma^2_{\nu} \sigma^2_{\nu}}$. Using $\pi_2(L) = \beta \hat{\pi}_2(L)$, one can use the resulting expression together with $\pi_1(L)$ to substitute in (A.56) and obtain (4.20) in Proposition 1. To complete the proof we need to argue why $\lambda(L) + \beta \tau_2 \frac{\lambda(L)}{L - \beta} = 0$ must have no solution inside the unit circle. Note that, from (A.62), the determinant of the matrix $\Xi^\star(L)$ is proportional to $\frac{\pi_1(L)}{\pi_1(0)}$ and, for the matrix to be a canonical factorization, the determinant must not vanish inside the unit circle. Because $\frac{\pi_1(L)}{\pi_2(L)} = \lambda(L) + \beta \tau_2 \frac{\lambda(L) - \lambda(0)}{L - \beta}$, for $\Xi^\star(L)$ to be the appropriate factorization, the right hand side must not vanish inside the unit circle.

**A.4 Proof of Proposition 2**

The candidate solution for the equilibrium in Proposition 2 is specified as

$$y_t = \tilde{Q}_z(L) \epsilon_t + \tilde{Q}_q(L) \eta_t. \quad (A.70)$$

The first part of the proof is equivalent to that of Proposition 1, up until equation (A.60). We then need to conjecture confounding dynamics, which we do by assuming that there exists a $\lambda \in (-1, 0)$ such that

$$\frac{\pi_1(\lambda)}{1 - \pi_2(\lambda)} = 0. \quad (A.71)$$

If this is the case then the matrix $\Xi^\star(z)$ in (A.62) harbors confounding dynamics since its determinant vanishes at $\lambda$. For notational convenience we assume that $\frac{\pi_1(\lambda)}{\pi_2(\lambda)} = \pi(L)(L - \lambda)$, which embeds conjecture (A.71). To obtain the canonical factorization of $g_s(z)$ we apply the steps in Appendix C.3 to $\Xi^\star(z)$ and we obtain

$$\Xi^\star(z) = \frac{1}{\sqrt{\sigma^2_{\pi} + \sigma^2_{\epsilon}}} \begin{pmatrix} \sigma^2_{\pi} + \sigma^2_{\epsilon} & 0 \\ \sigma^2_{\pi}(z - \lambda) \pi(z) & \sigma^2_{\pi}(1 - \lambda z) \end{pmatrix}. \quad (A.72)$$
For convenience define \( \tilde{\pi}_2(z) = (z - \lambda)\pi_2(z) \), and using \( \Xi^* (z) \) in the Wiener-Kolmogorov formula (A.61) one obtains the following two fixed point conditions in \( \pi(L) \) and \( \tilde{\pi}_2(L) \) after some straightforward rearrangements:

\[
\pi(L) \left[ (L - \lambda)(L - \zeta \tau) - (1 - \tau)\tilde{\pi}_2(L) \right] = \zeta \tau \lambda \pi(0) + A(L)L, \tag{A.73}
\]

and

\[
\frac{\tilde{\pi}_2(L)}{L - \lambda} = \frac{\zeta \tau (1 - \lambda\lambda)h_1(L) - (L - \zeta \tau)\frac{\sigma^2}{\sigma^2 \sigma^2} h_2(L)}{(L - \zeta (1 - \tau))(1 - \lambda L)h_1(L) - \zeta \frac{\sigma^2}{\sigma^2 \sigma^2} h_2(L)}, \tag{A.74}
\]

where

\[
h_1(L) \equiv \zeta \tau \lambda \pi(0) + A(L)L, \tag{A.75}
\]

and

\[
h_2(L) \equiv \pi(0) (L - \lambda) \left( \tau_0 - \tilde{\pi}(0) \frac{\sigma^2}{\lambda} \right) - \frac{\sigma^2}{(1 - \tau)} \frac{1 - \lambda^2}{\lambda} \left( A(\lambda) + \zeta \tau \pi(0) \right). \tag{A.76}
\]

Substituting (A.74) into (A.73) one obtains

\[
\pi(L)(L - \lambda) = \frac{(L - \zeta (1 - \tau))(1 - \lambda L)h_1(L) - \zeta \frac{\sigma^2}{\sigma^2 \sigma^2} Lh_2(L)}{(1 - \lambda L)(L - \zeta) (L - \zeta \tau (1 - \tau))}. \tag{A.77}
\]

We require \( \pi(L)(L - \lambda) \) to be stationary, which means that the two unstable roots at the denominator, \( \zeta < 1 \) and \( \zeta \tau (1 - \tau) < 1 \), need to be removed. In addition, our conjecture of confounding dynamics requires the left hand side expression to vanish at \( L = \lambda \). We can achieve all this by the appropriate choice of constants \( \pi(0), \lambda_2(0) \) and \( \lambda \). We thus have the following three conditions in three unknowns,

\[
(L - \zeta (1 - \tau))(1 - \lambda^2)h_1(\lambda) - \zeta \frac{\sigma^2}{\sigma^2 \sigma^2} \lambda h_2(\lambda) = 0, \tag{A.78}
\]

\[
\sigma^2(1 - \lambda \zeta)h_1(\zeta) - \zeta h_2(\zeta) = 0, \tag{A.79}
\]

\[
(1 - \zeta \tau (1 - \tau) \lambda)h_1(\zeta \tau (1 - \tau) + \zeta \frac{\sigma^2}{\sigma^2 \sigma^2} = 0. \tag{A.80}
\]

We first note that \( \lambda h_2(\lambda) = \tau_0 h_1(\lambda)(1 - \lambda^2) \), which implies that condition (A.78) is satisfied when \( \pi(0) = \frac{\lambda A(\lambda)}{\zeta \tau} \). Substituting this into the expressions for \( h_1(L) \) and \( h_2(L) \) one sees that,

\[
h_1(L) = LA(L) - \lambda A(\lambda), \quad \text{and} \quad h_2(L) = \frac{A(\lambda)}{\zeta \tau} (L - \lambda) \left( \tau_0 - \tilde{\pi}(0) \frac{\sigma^2}{\lambda} \right). \tag{A.81}
\]

Using these expressions into (A.79) one obtains

\[
\left( \tau_0 - \tilde{\pi}(0) \frac{\sigma^2}{\lambda} \right) = \tau \sigma^2 \frac{(1 - \lambda \zeta)}{A(\lambda)(\zeta - \lambda)} (\lambda A(\lambda) - \zeta A(\zeta)). \tag{A.82}
\]

Next define

\[
\tilde{\lambda}(L) \equiv (1 - \lambda L) \frac{LA(L) - \lambda A(\lambda)}{L - \lambda}. \tag{A.83}
\]
and note that condition (A.80) is satisfied when
\[ \dot{\lambda}(\zeta(1 - \tau_t)) + \dot{\lambda}(\zeta) \frac{\sigma_1^2}{\sigma_2^2 \sigma_3^2} = 0, \]  
(A.84)
which corresponds to (4.25) once we multiply both sides by \( \sigma_3^2/\sigma_2^2 \). With some additional straightforward algebra is then possible to solve for \( \pi(L) \) and \( \pi_2(L) \), and using the conditions, \( \dot{Q}_t(L) = \pi(L) \), and \( \dot{Q}_t(L) = \pi_1(L) \frac{\sigma_1^2(L)}{1 - \sigma_2(L)} \), equation (4.24) obtains.

A.5 Proof of Corollary 1 The proof of the corollary is a straightforward application of the following lemma.

**Lemma A4.** Consider the Real Business Cycle model (5.3)-(5.4). Let the information sets be specified as in (5.10). There exists a Rational Expectations Equilibrium with Confounding Dynamics of the form, \( k_{t+1} = K(L)\varepsilon_t \), and \( r_t = R(L)\varepsilon_t \), with
\[ K(L) = K(L) - (1 - \tau(\lambda))(1 - \lambda^2) \frac{A_k(\lambda)}{(1 - \lambda L)(\zeta - L)}, \]  
(A.85)
and, \( R(L) = A(L) - (1 - \alpha)K(L)L \), if there exists a \( \lambda \in (-1, 1) \), that solves
\[ R(\lambda)(\lambda - \zeta) = (1 - \alpha)(1 - \tau(\lambda))A_k(\lambda)\lambda, \]  
(A.86)
where \( K(L) \) and \( R(L) \) are the full information solutions, \( \tau(\lambda) \equiv \frac{A(\lambda)^2 \sigma_2^2}{A(\lambda)^2 \sigma_2^2 + \sigma_3^2} \). \( A_k(\lambda) \) is a function of \( \lambda \) that depends only on exogenous parameters, and \( R(L) \) has no zeros inside the unit circle other than \( \lambda \). If there exists only one \( \lambda \in (-1, 1) \) solving (A.86), the equilibrium is unique.

**Proof.** The proof follows the same steps as that of Theorem 1, with the difference that we solve for \( X(L) \) first – \( K(L) \) in the application. Recall that
\[ \phi_x = \alpha \beta, \quad \phi_y = 1 - \alpha \beta, \quad \phi_0 = 1, \quad \psi_x(L) = \alpha(1 + \beta) - \alpha L, \quad \psi_y(L) = 0, \quad \psi_0(L) = -1. \]
and
\[ \gamma_x(L) = (1 - \alpha)L, \quad \gamma_y(L) = 1, \quad \gamma_0(L) = -1. \]
Note that, although the notation adopted in the model has the two variables having different time subscripts, \( r_t \) and \( k_{t+1} \), they are both pre-determined at time \( t \), and so they are both functions of possibly the infinite history of \( \varepsilon_t \) up to time \( t \). Since we are looking for an equilibrium with confounding dynamics, we operationalize the condition \( R(\lambda) = 0 \) by conjecturing
\[ R(L) = (L - \lambda)G(L), \]  
(A.87)
where \( G(L) \) has no zeros inside the unit circle. Because in equilibrium \( R(L) = A(L) - (1 - \alpha)K(L)L \), the conjecture immediately implies
\[ A(\lambda) = (1 - \alpha)K(\lambda)\lambda, \]  
(A.88)
a relationship that will be useful in what follows. One important remark on (A.88) is that it implies \( \lambda \neq 0 \). In fact, evaluating the expression at \( \lambda = 0 \), provided that \( K(0) \) is well defined, which must be the case in the solution we want to characterize, gives \( A(0) = 0 \), which never holds by assumption. Hence, the statement of the Proposition requires \( |\lambda| \in (0, 1) \). The information set takes the form of (A.18), where \( x_{it} = a_{it} \) and \( y_t = r_t \), so that \( E_{it}(a_{it+1}) \) and \( E_{it}(r_{t+1}) \) are provided by (A.24) and (A.25), respectively. For the term \( E_{it}(k_{it+2}) \) things require some extra steps. We work under the conjecture that
\[ k_{it+1} = K(L)\varepsilon_t + V(L)\varepsilon_{it}, \]  
(A.89)
Next, we evaluate the variance-covariance generating function between the information set and \( k_{it+1} \), which is
\[ g_{k_{it}(a_{i,r})}(z) = \left[ K(z)A(z^{-1})\sigma_2^2 + V(z)\sigma_0^2 - K(z)(z^{-1} - \lambda)G(z^{-1})\sigma_3^2 \right]. \]  
(A.90)
We then use this expression, together with the canonical factorization \( \Gamma^*(z) \) in (A.20) in the Wiener-Kolmogorov
formula (2.4), and following steps similar to (A.28) and (A.29) to finally get

\[ \mathbb{E}_{it}(k_{it+2}) = L^{-1} \left[ \begin{array}{c} K(L) - K(0) \\ (1 - \tau(\lambda)) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \end{array} \right] \varepsilon_t - \frac{1}{\lambda(1 - \lambda L)} \left[ K(0) - K(\lambda) - (V(0) - V(\lambda))A(\lambda) \right] \varepsilon_t \\
+ L^{-1} \left[ V(L) - V(0) \right] v_{it} - \tau(\lambda) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ \frac{K(0) - K(\lambda)}{\lambda(1 - \lambda L)} + (V(0) - V(\lambda)) \right] v_{it} \]

(A.91)

We can now use the expressions for the expectational terms to obtain a fixed point condition similar to (A.32),

\[
\alpha(1 + \beta)K(L)\varepsilon_t + \alpha(1 + \beta)V(L)v_{it} = \alpha\beta L^{-1} \left[ K(L) - K(0) \right] \varepsilon_t + \alpha\beta L^{-1} \left[ V(L) - V(0) \right] v_{it} \\
- \alpha\beta(1 - \tau(\lambda)) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ K(L) - K(0) - (V(L) - V(0))A(\lambda) \right] \varepsilon_t + \alpha\beta \frac{\tau(\lambda)}{\lambda(1 - \lambda L)} \left[ K(L) - K(0) - (V(L) - V(0))A(\lambda) \right] v_{it} \\
+ \alpha K(L)\varepsilon_t + \alpha V(L)Lv_{it} + A(L)\varepsilon_t + v_{it} - A(L) - A(0) \right] L^{-1} \varepsilon_t + (1 - \tau(\lambda)) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ A(\lambda) - A(0) \right] \varepsilon_t \\
- \frac{\tau(\lambda)}{\lambda(1 - \lambda L)} \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ A(\lambda) - A(0) \right] v_{it} + (1 - \alpha\beta) \left[ A(L) - (1 - \alpha)K(L)L - A(0) \right] \varepsilon_t + (1 - \alpha\beta)(1 - \tau(\lambda)) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ A(\lambda) - A(0) \right] v_{it},
\]

(A.92)

where we have used \((L - \lambda)G(L) = A(L) - (1 - \alpha)K(L)L\), and thus \(-\lambda G(0) = A(0)\), to substitute for terms related to \(G(L)\). The fixed point equation contains only terms related to the endogenous polynomials \(V(L)\) and \(K(L)\), and one can proceed to solve for the fixed point as in the proof of Theorem 1. In particular, using the same steps as in Lemma A2, one can show that \(A(\lambda)V(\lambda) = K(\lambda)\), and, in addition, we know that (A.88) holds, so we can set \(K(\lambda) = \frac{A(\lambda)}{\lambda(1 - \alpha)}\). The uniqueness of a stationary solution under Assumption (s) and condition (5.7), is once again obtained by the appropriate choice of \(V(0)\) and \(K(0)\). In the end, the expression for \(A_k(\lambda)\), analogue to the constant \(A(\lambda)\) in Theorem 1, can be simplified to

\[
A_k(\lambda) = \frac{(1 - \lambda\beta)(1 - \lambda^2)\beta(1 - \beta)(\eta(1 - \beta\alpha) - 1))}{\lambda(1 - \lambda\beta)(\zeta - \lambda)(1 - \lambda^2)(\zeta - \beta)} \left( 1 + \tau(\lambda) \right) A(0) \right] L^{-1} \varepsilon_t + (1 - \alpha\beta)(1 - \tau(\lambda)) \frac{1 - \lambda^2}{\lambda(1 - \lambda L)} \left[ A(\lambda) - A(0) \right] v_{it},
\]

(A.93)

The condition for the existence of one \(|\lambda| \in (0, 1)\) follows from using \(K(\lambda)\) to write \(R(L)\) and then imposing \(R(\lambda) = 0\). The same argument that we have used in the proof of Theorem 1 to argue that when the equilibrium coefficients are evaluated using the \(\lambda\) that solves \(R(\lambda) = 0\), there must be no other points at which \(R(L)\) vanishes inside the unit circle, applies here as well. This completes the proof.

The proof of Corollary 1 consists in plugging \(A(L) = \frac{1 + \beta L}{1 - \rho L}\) into the above expressions and rearranging terms when possible.