Firestorm: Multiplicity in Models with Full Information *

Jonathan J Adams[†]

October 31, 2022

WORKING PAPER

Link to Most Current Version

Abstract

Dynamic stochastic models with full information and rational expectations (FIRE) are not as well determined as is commonly believed. If the assumption of causality is relaxed so that prices and decisions may anticipate future shocks, then FIRE models generally feature multiple equilibria. The multiplicity is due to the endogenous feedback from choices to information to choices, which in equilibrium may contain self-fulfilling news about future shocks. I demonstrate the multiplicity in several examples, including canonical asset pricing and business cycle models. To motivate relaxing the causality assumption, I also study examples with apparent non-causality, even if the model is fundamentally causal. Then I examine how the multiplicity arises in a dynamic programming problem with decentralized markets.

JEL-Codes: C62, D50, D84, E32

Keywords: News, Self-Fulfilling Equilibria, Multiplicity, Rational Expectations

---

*I am grateful for many helpful comments, especially from Philip Barrett, Martin Beraja, Ryan Chahrour, Kris Nimark, Iván Werning, and Stephen Wright. All errors are my own.

[†]University of Florida. Website: www.jonathanjadams.com Email: adamsjonathan@ufl.edu

1 Introduction

Mainstream dynamic stochastic models have more equilibria than previously understood. Theories typically assume that equilibria must be functions of current and past shocks. But if the causality assumption is relaxed, then mainstream models generally feature multiple equilibria. This multiplicity is due to the full information assumption: if agents in the model can observe endogenous variables, then their information sets are endogenous too. The endogenous variables and information sets are not jointly determined by the assumptions of a typical full information rational expectations (FIRE) model. Instead, equilibria feature *self-fulfilling news*.

The causality assumption is overly restrictive. In business cycle models, equilibria are functions of stochastic shocks. But it is not obvious that equilibria must be *causal* functions of these shocks. Indeed, a large empirical literature studying news in the macroeconomy concludes that at least a portion of aggregate shock processes are anticipated.[1] This is unsurprising, given that common shocks such as "productivity" or "demand" are not fundamentally unpredictable. Rather, productivity and demand shocks are functions of many underlying processes, and information about the underlying processes can appear to be non-causally related to the aggregate shocks.[2] Moreover, such information can arise endogenously, manifesting as self-fulfilling news about future shocks.

How can news be self-fulfilling? Agents' information sets include everything they observe. The full information assumption implies that this includes both fundamental shocks and endogenous macroeconomic variables. But endogenous variables are determined by forward-looking decisions, so they depend on information sets. This is a feedback mechanism: the information set determines the information set! Is there a unique solution to this system? Not in general in FIRE models.

What of the existing uniqueness theorems? Macroeconomists rely on well-known regularity conditions (e.g. Blanchard and Kahn (1980)) to ensure that a model has

[1]Papers such as Cochrane (1994), Beaudry and Portier (2006), and Beaudry and Lucke (2010) document evidence for large business cycle effects of news using vector autoregressions. Beaudry and Portier (2014) survey the news literature at large.

[2]Many empirical papers relax the "fundamentalness" assumption that shocks must be causally invertible from observed time series; examples include Lippi and Reichlin (1994), Mertens and Ravn (2010), Lanne and Saikkonen (2013), Forni, Gambetti, Lippi, and Sala (2017) and Chahrour and Jurado (2021).

a unique equilibrium, but crucially, this uniqueness is *conditional on the process for information*. The Blanchard-Kahn condition or its generalizations do not guarantee that any joint determination of information and decisions has a unique solution.

The multiplicity is pervasive. In a simple asset pricing model, I show that in general there are multiple equilibria that are consistent with the full information assumption. I explore the multiplicity with several numerical examples. First I demonstrate how to construct valid equilibria from non-causal information sets. Then I consider several mechanisms to motivate non-causality, including a model with non-invertible stochastic processes, a model where agents receive exogenous news about the future, and a model of complexity where the shocks process is determined by an underlying unobservable multivariate process. Next, I move on to a general business cycle model, prove that it features multiplicity, and numerically demonstrate multiple equilibria in the canonical RBC model.

Are equilibria with self-fulfilling news just a mathematical curiosity that can be safely disregarded? I argue that they are not: equilibria with self-fulfilling news are observationally equivalent to traditional equilibria given appropriate initial conditions for the information set. So there is no way to uniquely select one equilibrium without self-fulfilling news. In other words, even if equilibria with self-fulfilling news are assumed away, the potential for multiplicity remains, with the economy's long run dynamics determined by the model's initial conditions. This is because the same information feedback that generates self-fulfilling news also generates self-perpetuating news.

The implication is a critique of FIRE in macroeconomics: *full information rational expectations models are incomplete*. Economic objects such as prices and quantities must be determined jointly with information processes, because they are both endogenous. And full information models do not determine them uniquely. Fortunately, tools to resolve this incompleteness already exist. The long literature relaxing FIRE in macroeconomics includes many theories in which information and actions are jointly determined in equilibrium, including the seminal Lucas (1972). If full information fails, then information frictions or bounded rationality are necessary.

The multiplicity in this paper is most closely related to a set of papers in the information frictions literature that feature agents learning from endogenous signals, joint determination of information sets and actions, and multiple equilibria. Macroeconomic examples include Benhabib, Wang, and Wen (2015), Gaballo (2018), and Chahrour and

Gaballo (2020) among others. Thematically related is Morris and Shin (1998) which studies global games with multiple equilibria, such as bank runs or currency crises, and shows that abandoning full information resolves the multiplicity. However, while Morris and Shin's resolution to the multiplicity is the same (relaxing full information), the source of their multiplicity is entirely different, stemming from strategic interaction in the global games. The models I study feature no such strategic interaction; rather, the multiplicity is due entirely to the information endogeneity.

The remainder of the paper is organized as follows. Section 2 describes self-fulfilling news in a simple two-period model, to demonstrate most clearly how multiplicity can arise mathematically, although the additional equilibrium is unintuitive and unrealistic. Section 3 explores the multiplicity in an asset pricing model with many periods, where some concerns from the two-period model do not apply. Section 4 argues that self-fulfilling equilibria cannot be assumed away. Section 5 motivates the possibility of non-causality, and shows how multiplicity is possible even when the model is fundamentally causal. Section 6 demonstrates how the multiplicity arises in general dynamic macroeconomic models. Section 7 concludes.

2 An Asset Pricing Model With Two Periods

I introduce the concept of self-fulfilling news in a two period asset pricing model. Relaxing the assumption that prices must be causal allows for perfect foresight to be a valid equilibrium. But even though it satisfies the technical requirements, the perfect foresight equilibrium is disregardable for a number of reasons.

Consider a standard two-period asset pricing model. In the first period, agents trade an asset at price p_1 which pays stochastic dividend $x_2 \sim N(0,1)$ in the second period. Agents with information set Ω_1 price the asset by

$$p_1 = E[x_2|\Omega_1] \tag{1}$$

Agents have *full information*, so their information set includes the price:

$$\Omega_1 = p_1 \tag{2}$$

Definition 1 defines equilibrium for this two period asset pricing model:

Definition 1 *A* ***full information equilibrium of the two-period asset pricing model*** *is a price p_1 and information set Ω_1, given a stochastic dividend x_2 such that:*

1. *The price* p_1 *satisfies the asset pricing equation* (1)

2. *The information set is given by* (2)

If the price must be a causal function of dividends, then this model has only one equilibrium, which I call the *nescient* equilibrium. The nescient price p_1^N is

$$p_1^N = E[x_2|\Omega_1^N] = 0$$

which is rationalized by the nescient information set $\Omega_1^N = p_1^N = 0$. This is the standard equilibrium: agents have no information about the future, so they forecast by the unconditional expectation.

But if the price is allowed to be non-causal, potentially containing information about the future, then there is also a *non-nescient* equilibrium.[3] This non-nescient price p_1^{PF} is

$$p_1^{PF} = x_2$$

This price perfectly predicts the dividend. It satisfies the pricing equation (1):

$$p_1^{PF} = E[x_2|\Omega_1^{PF}] = x_2$$

because the non-nescient information set is just the future dividend:

$$\Omega_1^{PF} = p_1^{PF} = x_2$$

The non-nescient equilibrium $\{p_1^{PF}, \Omega_1^{PF}\}$ is plainly absurd. It satisfies the technical definition of an equilibrium, but can usually be ignored. It has several fundamental problems, each of which might justify assuming it away:

1. The non-nescient equilibrium features perfect foresight, which may be unrealistic and defeats the purpose of considering a stochastic problem at all.

2. The non-nescient equilibrium is not robust to delayed observation of the price: agents must both see the price and value the asset simultaneously. What if the market is a sealed bid auction?

[3] The perfect foresight equilibrium discussed here is the unique non-nescient equilibrium that is linear in the exogenous dividends. There may exist other non-nescient equilibria that are nonlinear or depend on extrinsic noise.

3. Nor is the non-nescient equilibrium robust to noisy observation of the price: if traders observe the equilibrium prices with even a trivial amount of noise when determining their expectations, the non-nescient equilibrium disappears.

These are each fundamental problems for the non-nescient equilibrium; it can reasonably be disregarded.

Is this conclusion – that non-nescient equilibria can be ignored – robust to generalizing the model? No: when there are many time periods, each of these concerns no longer holds. There is a continuum of non-nescient equilibria beyond the perfect foresight case, these equilibria are still valid even when the price is observed with a delay or with noise, and agents know with certainty which equilibrium they are in. The next section explores this generalization.

3 An Asset Pricing Model With Many Periods

This section introduces the multiplicity in a simple asset pricing model. I prove a general condition for multiplicity of stationary equilibria. I show that each stationary equilibrium is associated with an information basis. Then I calculate equilibria in several examples. Finally, I show how criteria that eliminate non-nescient equilibria in the two-period model do not apply to the infinite case.

3.1 Full Information Equilibria

A standard linear asset pricing model[4] is characterized by three equations:

$$p_t = x_t + \beta E[p_{t+1}|\Omega_t] \tag{3}$$

$$x_t = \sum_{j=0}^{\infty} X_j \varepsilon_{t-j} \tag{4}$$

$$\Omega_t = \{\Omega_{t-1}, \varepsilon_t, p_t\} \tag{5}$$

[4]This asset pricing model can represent a variety of settings. p_t may be the price of an equity after a dividend x_t is announced but before it is paid, with risk-neutral traders. p_t may be the utility-denominated price of an asset paying utility dividend x_t as in Lucas (1978). p_t may be the equilibrium price of an asset with stochastic supply x_t where agents have CARA preferences, which is common in the finance literature, and appears in settings with learning from endogenous information sets, such as Singleton (1987) and Nimark (2017). Or the asset pricing equation may be the linearized form of a more complicated model.

Equation (3) is the asset pricing equation, whereby the price p_t of an asset is determined by a stochastic dividend x_t, and the discounted expectation of the next period's price, conditional on the information set Ω_t and with discount factor $\beta \in (0, 1)$. Equation (4) gives the time series process for the dividend x_t as a moving average of white noise shocks $\varepsilon_t \sim N(0, 1)$. The coefficients $\{X_j\}_{j=0}^{\infty}$ determine the time series properties of x_t.

The *full information* assumption manifests as equation (5), which describes how the information set Ω_t evolves. At time t, agents observe the current fundamental shock ε_t, the current asset price p_t, as well as everything they knew in the previous period Ω_{t-1}. Thus their information set is the entire history of ε_t's and p_t's. Crucially, ε_t is exogenous, but p_t is endogenous.

Formally, an equilibrium of this asset pricing model is given by Definition 2:

Definition 2 *A **full information equilibrium of the asset pricing model** is a sequence of prices p_t, dividends x_t, and information sets Ω_t, given a sequence of shocks ε_t such that for all t:*

1. *Prices satisfy the asset pricing equation* (3)

2. *The dividend evolves by equation* (4)

3. *The information set evolves by equation* (5)

Furthermore, the equilibrium is called *stationary* if the price p_t and dividend x_t processes have a time-invariant autocovariance function.

This is a standard setting. How can there be many full information equilibria? The multiplicity arises from the joint determination of endogenous prices and the information set. For any information set Ω_t, the equilibrium price is uniquely determined as the discounted sum of future dividends:

$$p_t = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t\right] \tag{6}$$

However, the information set also depends on the price, by the full information assumption (5). So the price p_t appears on both sides of the equilibrium equation (6):

$$p_t = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \{\Omega_{t-1}, \varepsilon_t, p_t\}\right] \tag{7}$$

The pricing equation is self-referential! Without conditioning on Ω_t, the equilibrium price may no longer be determined. Of course, perhaps (7) has a unique fixed point for p_t? Theorem 1 says no: there are many equilibrium prices that satisfy (7), except under specific conditions.

Before demonstrating the multiplicity, it is necessary to introduce notation for two special equilibria. The first special equilibrium is the *nescient* equilibrium, which is the familiar equilibrium that economists typically select when assuming full information:

Definition 3 *A full information equilibrium is called the* ***nescient*** *equilibrium if the equilibrium information set is the nescient information set* Ω_t^N*, given by:*

$$\Omega_t^N = \{\varepsilon_{t-j}\}_{j=0}^{\infty} \tag{8}$$

The nescient equilibrium is the case where prices reveal no information about future shocks, so expectations depend only on the history of fundamental shocks. Applying equation (6), the nescient price is given by

$$p_t^N = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t^N\right] \tag{9}$$

$$= E\left[\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \beta^j X_k \varepsilon_{t+j-k} | \Omega_t^N\right] = \sum_{j=0}^{\infty}\sum_{k=j}^{\infty} \beta^j X_k \varepsilon_{t+j-k}$$

The second special equilibrium is perfect foresight:

Definition 4 *A full information equilibrium is called the* ***perfect foresight*** *equilibrium if the equilibrium information set is the perfect foresight information set* Ω_t^{PF}*, given by:*

$$\Omega_t^{PF} = \{\varepsilon_{t-j}\}_{j=-\infty}^{\infty} \tag{10}$$

The perfect foresight equilibrium is the case where agents know the values of all future and past shocks. Applying equation (6), the perfect foresight price is given by

$$p_t^{PF} = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t^{PF}\right] \tag{11}$$

$$= \sum_{j=0}^{\infty} \beta^j x_{t+j}$$

If the nescient and perfect foresight prices differ, then it is possible for the asset price to contain some news about future dividends. Agents then use that news to price the asset. *News from prices begets news from prices.* Theorem 1 formalizes the possibility of this self-fulfilling news.

Theorem 1 *If $p_t^N \neq p_t^{PF}$ for some t, then there exist multiple stationary full information equilibria of the linear asset pricing model*

Proof: Appendix A.1

The intuition for the proof is that the price contains everything in an information set that is relevant for pricing an asset, so if agents see prices generated by any information set, they will price the asset as if they possessed the entire information set. The perfect foresight price p_t^{PF} does exactly this, and is always a valid equilibrium by assumption, so there must be multiplicity. Of course, agents trading in this equilibrium do not literally have perfect foresight – observing p_t^{PF} does not tell them the future path of shocks – rather, they observe a time series of prices that forecasts discounted future dividends better than they could have using past shocks alone.

3.2 Information Bases

Stationary equilibria are entirely characterized by an *information basis*, which I study in this section.

3.2.1 Notation for Stationary Equilibria

When describing stationary equilibria, it is convenient to use lag operator notation. For example, equation (4) implies that the dividend time series x_t is characterized by the lag operator polynomial $X(L) = \sum_{j=0}^{\infty} X_j L^j$:

$$x_t = X(L)\varepsilon_t$$

The summation starts at zero because x_t is a causal process. I will write similar polynomials for other time series. However, other lag operator polynomials may be non-causal. Additionally, for some process $W(L) = \sum_{j=-\infty}^{\infty} W_j L^j$, let $W^*(L)$ denote the *adjoint* of the polynomial:

$$W^*(L) = \sum_{j=-\infty}^{\infty} W'_{-j} L^j$$

which is just $W(L)$ with the coefficients reversed and transposed.

3.2.2 Information Basis Definition and Properties

A stationary equilibrium is characterized by its *information basis*. The information basis is a white noise process w_t that captures all of the new information acquired by agents at time t. It is called a basis because all equilibrium time series can be expressed as a linear combination of w_t's. In the simple asset pricing model, w_t is the white noise component of the Wold representation of the equilibrium price process. Definition 5 gives the properties of a valid information basis.

Definition 5 *A process w_t is an* ***information basis*** *for a stationary equilibrium price process p_t if:*

1. *ε_t is causally spanned by w_t, i.e. there exists a lag operator polynomial $U(L)$ such that $\varepsilon_t = \sum_{j=0}^{\infty} U_j L^j w_t$*

2. *w_t is a process of innovations for the Wold representation of $P(L)$, i.e. for $p_t = P(L)w_t$, $P(L)$ is both causal and causally invertible, while w_t is unit white noise so that*

$$cov(w_t, w_{t-j}) = \begin{cases} 1 & j = 0 \\ 0 & j \neq 0 \end{cases}$$

Why must a valid information basis have the properties in Definition 5? First, agents know the entire history of current and past shocks, so the current shock ε_t must be recoverable from current and past information innovations w_t. Next, the only way agents can learn information beyond that contained in fundamental shocks is from current and past prices, so the basis must be a causal linear combination of prices and their Wold innovations. Finally, there are many ways to write a time series basis that is causal in the Wold innovations of the price process, and the Wold innovations themselves are a convenient choice. Given that the information basis causally spans both shocks and prices, the information set (5) can be reduced to the infinite history of the w_t's:

$$\Omega_t = \{w_{t-j}\}_{j=0}^{\infty} \tag{12}$$

This form of the Wold representation is nonstandard. In practice, economists often normalize the process of Wold innovations so that they represent forecast errors. However, other normalizations are valid too; in this case I normalize the innovations to have unit variance. This improves notation in many cases, because their associated

Blaschke factors are simpler, and because a unit variance innovation is a unitary transformation of the fundamental shock space.[5] This is a consequence of Theorem 2, which says that the polynomial $W(L)$ characterizing a basis must be anti-causal and unitary, i.e. causally invertible by its adjoint:

Theorem 2 *Any information basis w_t is given by $w_t = W(L)\varepsilon_t$ with causal inverse*

$$W^{-1}(L) = W^*(L)$$

Proof: Appendix A.2

The equilibrium price process is uniquely determined from the information basis, by Theorem 3. Some notation is needed first. Let the lag operator polynomial $P(L)$ denotes the process for the price in terms of the information basis, so that $p_t = P(L)w_t$. And let $W(L)$ denote the process for the information basis in terms of the fundamental shocks, so that $w_t = W(L)\varepsilon_t$. The inverse of $W(L)$ is its adjoint, so shocks can be written as $\varepsilon_t = W^*(L)w_t$. Theorem 2 implies that $W^*(L)$ is strictly causal. Lastly, let $[\cdot]_+$ denote the annihilation operator, which annihilates strictly negative powers of L.

Theorem 3 *For any information basis w_t that spans the fundamental shocks by $\varepsilon_t = W^*(L)w_t$, the equilibrium price process is given by $p_t = P(L)w_t$, where:*

$$P(L) = [(I - \beta L^{-1})^{-1} X(L) W^*(L)]_+ \tag{13}$$

Proof: Appendix A.3.

The polynomial for the information basis $W(L)$ determines everything about a particular equilibrium. What might $W(L)$ look like?

The space of potential information basis polynomials $W(L)$ is restricted by the requirements that it be unit white noise and that its adjoint $W^*(L)$ is causal. A large class of candidates is the set of non-causal Blaschke products,[6] so that the basis is given by

$$W(L) = \prod_{k=1}^{K} B_k \tag{14}$$

[5] This property holds in the asset pricing setting where prices and shocks are scalar-valued. Information bases become more complicated when there is more than one shock or endogenous variable, as in Section 6.

[6] Few other candidates exist. The coefficients of $W(L)$ represent the Taylor coefficients of an analytic inner function, which must be a product of Blaschke factors and singular inner functions, which are not explored in this paper. Such products need not have finite order.

where

$$B_k = (\phi_k - L^{-1})(1 - \phi_k L^{-1})^{-1} \qquad |\phi_k| < 1$$

In this case, $W(L)$ is called a *non-causal Blaschke product of order* K. Still, even first order Blaschke products are general enough to contain many interesting cases; Theorem 4 gives three examples.

Theorem 4 *A first order Blaschke information basis given by*

$$w_t = (\phi - L^{-1})(I - \phi L^{-1})^{-1}\varepsilon_t$$

for $|\phi| < 1$ *contains three special cases:*

1. *When* $\phi = 1$ *the equilibrium price* p_t *is the nescient equilibrium*
2. *When* $\phi = 0$ *the equilibrium price* p_t *is the equilibrium for one-period-ahead news*
3. *When* $\phi = \beta$ *the equilibrium price* p_t *is the perfect foresight equilibrium*

Proof: Appendix A.4

A first-order Blaschke information basis can produce a variety of equilibria by varying a single parameter, the Blaschke root ϕ. The nescient and one-period-ahead news cases feature information bases of $w_t = \varepsilon_t$ and $w_t = \varepsilon_{t+1}$ respectively. But these familiar cases bookend a variety of intermediate bases that vary continuously with ϕ. Most surprising, $\phi = \beta$ yields the perfect foresight equilibrium; the corresponding information basis is the one that is exactly disentangled by the forward-looking price setting equation (11).

3.3 Example: AR(1) Dividend

I have established that the asset pricing model may exhibit many equilibria, depending on the information basis. What do these equilibria look like? This section explores the multiplicity with a simple example.

Suppose that the dividend is AR(1) with autocorrelation $\rho \in (0, 1)$:

$$x_t = \rho x_{t-1} + \varepsilon_t \tag{15}$$

The lag operator polynomial associated with this dividend is

$$X(L) = \sum_{j=0}^{\infty} \rho^j L^j$$

I solve for the equilibria associated with a range of information bases. In each case, the information basis $w_t = W(L)\varepsilon_t$ is given by a single non-causal Blaschke factor:

$$W(L) = (\phi - L^{-1})(1 - \phi L^{-1})^{-1} \tag{16}$$

which means that the information process can be written recursively by

$$w_t = \phi w_{t+1} + \phi \varepsilon_t - \varepsilon_{t+1}$$

I calculate the equilibria for 20 values of ϕ, equally spaced on $[0, 1]$. Per Theorem 4, these equilibria are bookended by the nescient case ($\phi = 1$) and one-period-ahead news ($\phi = 0$). These Blaschke factors satisfy most information basis properties from Definition 5 by construction; I numerically confirm that the implied $P(L)$ is the Wold representation of the price for all calculated equilibria.[7]

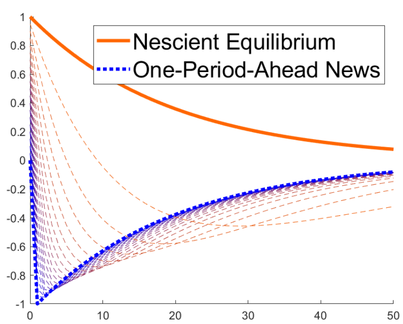
Figure 1 plots the impulse response functions (IRFs) associated with the different equilibria, with the two special cases highlighted for $\phi = 1$ and $\phi = 0$. In the intermediate cases, the color gradient indicates the position of the parameter ϕ on the unit interval. Lines of the same shade across panels correspond to the same information basis.

Panels (a) and (b) present causal IRFs, because they show responses to a unit innovation in the information w_t. Panel (a) plots the IRFs for the dividend x_t; these IRFs are given by the polynomial coefficients of $X(L)W^*(L)$. The IRFs associated with different ϕ's become negative because changing ϕ from 0 to 1 flips the sign on the shock, i.e. in the nescient case $w_t = \varepsilon_t$ but in the one-period-ahead news case $w_t = -\varepsilon_{t+1}$. But even though the IRFs change depending on ϕ, they all generate the exact same process for x_t, and all have the exact same autocovariance functions. Panel (b) plots the IRFs for prices, generated by the polynomials from Theorem 3. In effect, these are just the nescient asset prices for the corresponding dividends given in Panel (a). Thus when dividends have negative IRFs, the prices have negative IRFs. And when ϕ starts to decrease from 1, this transition occurs quickly but smoothly.

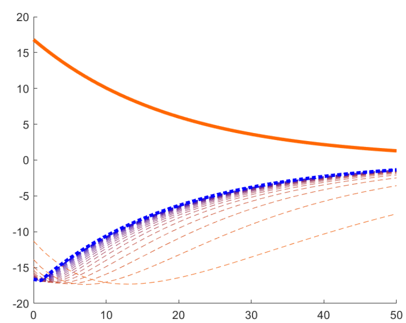
In contrast, Panels (c) and (d) present non-causal IRFs, because they show responses to a unit fundamental shock ε_t. Panel (c) plots the IRFs for the information

[7]This property may not hold in general, i.e. for some models, conjecturing an information basis may imply a process for prices with a Wold representation whose innovations form a new basis. This will be the case in the business cycle model in Section 6.2, at which point I discuss how to numerically confirm this property.

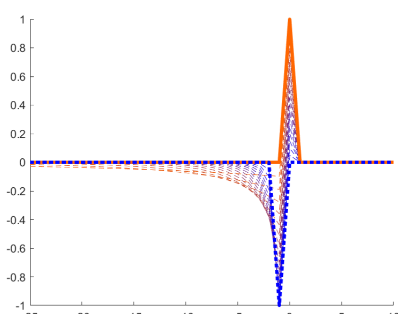
bases, which are given by the polynomial $W(L)$. These are just plots of Blaschke factors, which spike at 0 and -1. In the special cases, only these values are non-zero, but over the smooth transition from nescient to one-period-ahead news, the information basis contains a linear combination of many future shocks. There are no nonzero values to the right of the origin because an information innovation w_t must be orthogonal to past shocks. Panel (d) plots the response of prices to fundamental shocks. These IRFs are both causal (because dividends are affected by past shocks) and non-causal (because current information contains news about future shocks.) The IRFs are given by the polynomial $P(L)W(L)$, which is the convolution of the IRFs in Panel (c) with their pairs in Panel (b). In the nescient case, the price is unaffected by future shocks, while in the one-period-ahead news case, the price is only affected by the next period's shock.



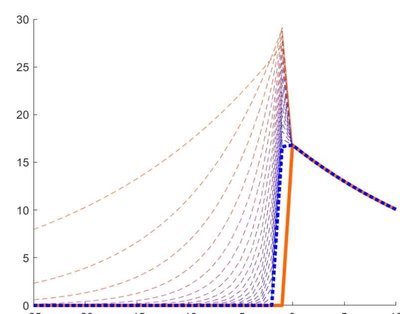
(a) Dividend IRF: Information Basis



(b) Price IRF: Information Basis



(c) Information IRF: Shock Basis



(d) Price IRF: Shock Basis

Figure 1: Multiple Equilibria in the AR(1) Example

Asset price volatility depends on the information basis. This is because giving agents news about future shocks necessarily increases the variance of their forecasts of future dividends. Figure 1 makes this sensitivity to the information basis clear: the price variance is the sum of the squares of the IRF points in Panel (d). The asset price has the lowest variance in the nescient equilibrium, and higher variance one agent's have one-period-ahead news. But intermediate cases increase the variance substantially, and achieve the upper bound on price volatility in the perfect foresight case when $\phi = \beta$.

Are non-causal IRFs problematic? Certainly not from an applied perspective; empirical researchers regularly estimate non-causal IRFs to find the effects of news about future fundamentals on current prices or other endogenous variables. Indeed, the estimated non-causal components are sometimes large (e.g. Beaudry and Portier (2006), Gazzani (2020), or especially Chahrour and Jurado (2021)).

Still, non-causality is one property that theorists might be tempted to axiomatically assume away in order to pin down equilibrium. This would be unwise; time series that appear non-causal in the data may have an underlying causal representation. Section 5 demonstrates that multiplicity can occur even when the information bases are entirely causal.

3.4 Imperfect Observation of the Price

Does the multiplicity exist when agents cannot perfectly observe the asset price? In this section I relax perfect observation in two ways: delayed observation and noisy observation. In the two-period model, either of these assumptions eliminates the non-nescient equilibrium. This is not so when time is infinite.

3.4.1 Delayed Observation of the Price

In this section, I assume that agents do not observe the contemporaneous price when forming their expectations. For example, this occurs when the price of an asset is determined by a sealed-bid auction.

In the two-period model of Section 2, this assumption eliminates all non-nescient equilibria. When prices are observed, the information set was $\Omega_1 = p_1$. But without observing prices, $\Omega_1 = \emptyset$. Thus the equilibrium price with delayed observation was given by $p_1 = E[x_2] = 0$, the nescient equilibrium.

But delayed observation does not eliminate non-nescient equilibria when time is infinite. The law of motion for the information set is now:

$$\Omega_t = \{\Omega_{t-1}, \varepsilon_t, p_{t-1}\} \tag{17}$$

so that agents only observe price p_{t-1} in period t. The remaining model assumptions are unchanged, so that equilibrium sequences must satisfy equations (3), (4), and (17).

Under this assumption, the multiplicity persists. This is because for any equilibrium sequence of prices p_t^{PO} for the model with perfect observation (i.e. satisfying Definition 2), observing ε_t and p_{t-1}^{PO} reveals p_t^{PO}. Theorem 5 formalizes this result.

Theorem 5 *If the price process $p_t^{PO} = P(L)w_t$ is an equilibrium price for the model with perfect observation, with information basis $w_t = W(L)\varepsilon_t$, then it is also an equilibrium price for the delayed observation model.*

Proof: Appendix A.5

The proof relies on the properties of information bases: they are invertible by their adjoint (Theorem 2), which is strictly causal. Therefore, past observations of the basis and the current fundamental shock jointly reveal the current basis. Thus the current price is inferable, even if it is not directly observed.

3.4.2 Noisy Observation of the Price

In this section, I assume that agents do not observe the equilibrium price directly. Instead, they observe a noisy signal s_t of the price:

$$s_t = p_t + \nu_t \tag{18}$$

where ν_t is exogenous white noise with $\nu_t \perp x_{t+j}$ for all j.

In the two-period model (Section 2) noisy observation eliminates the non-nescient equilibrium. Under this assumption, the information set is $\Omega_1 = s_1$, and the equilibrium price is given by

$$p_1 = E[x_2|s_1] = \frac{cov(x_2, s_1)}{var(s_1)} s_1 = \frac{cov(x_2, p_1)}{var(p_1) + var(\nu_1)}(p_1 + \nu_1)$$

whose only solution is $p_1 = 0$ (the nescient equilibrium) if $var(\nu_1) > 0$.

But when time is infinite, multiplicity prevails. The law of motion for the information set is now:

$$\Omega_t = \{\Omega_{t-1}, \varepsilon_t, s_t\} \tag{19}$$

so that agents forecast conditional on the current shock ε_t and the noisy signal s_t, rather than the price itself. Prices p_t are still formed by rational expectations per equation (3).

Any equilibrium price of the model with perfect observation p_t^{PO} is also an equilibrium of the model with noisy observation. This is because while the noisy signal does not reveal the price alone, it is able to do so when combined with the history of signals and dividends. Theorem 6 proves this result.

Theorem 6 *If the price process $p_t^{PO} = P(L)w_t$ is an equilibrium price for the model with perfect observation, with information basis $w_t = W(L)\varepsilon_t$, then it is also an equilibrium price for the noisy observation model.*

Proof: Appendix A.6

Theorem 6 looks similar to Theorem 5 but the proof takes an entirely different strategy: I conjecture that w_t and ν_t form a two-dimensional information basis, show that the basis can be recovered from observed signals, and then show that the perfect observation price p_t^{PO} is also the price implied by the conjectured information basis.

4 Axiomatic Nescience?

There is a multiplicity of equilibria, and rejecting perfect foresight is not enough to guarantee uniqueness. Is it feasible to axiomatically select the nescient equilibrium, eliminating self-fulfilling news and returning to uniqueness? In this section, I argue that it is not.

A continuum of equilibria in the baseline model can be rationalized as nescient equilibria of a model with a finite history, by choosing an appropriate initial condition. Thus a nescience axiom does not uniquely select an equilibrium. This is true even in the long run, where nescient equilibria feature a continuum of possible autocovariance functions. The same self-referential information feedback that generates self-fulfilling news in the infinite history model generates self-perpetuating dependence on initial conditions in the finite history model.

4.1 Finite History Model

I now modify the asset pricing model from Section 3 so that time no longer runs infinitely into the past. Instead, agents are endowed with a signal about future dividends as an initial condition of the model. Equilibrium prices, dividends, and information are determined as before, satisfying equations (3), (4), and (5), except time begins with $t = 1$.

However, a discrete beginning for time breaks the recursive structure of the information set evolution equation (5), so an initial condition is required. Therefore, the initial information set is given by Ω_0, which may contain information about future dividends. To clarify the structure of Ω_0, consider it as being determined by an exogenous signal process s_t that contains news about future shocks. Agents only receive a signal in the initial period, but it is useful to write it more generally, defined by

$$s_t = S(L^{-1})L^{-1}\varepsilon_t \tag{20}$$

where $S(L)$ is some causal lag operator polynomial, so that $S(L^{-1})$ reveals information about future shocks $L^{-1}\varepsilon_t$. This general structure allows for simple one-period-ahead news ($S(L) = 1$) but also recursive structures so that agents learn about the next period's dividend but also future news signals.

An equilibrium of this finite history asset pricing model is given by Definition 6:

Definition 6 *A* ***full information equilibrium of the finite history asset pricing model*** *is a sequence of prices p_t, dividends x_t, and information sets Ω_t, given a sequence of shocks ε_t and initial condition Ω_0 such that for all $t \geq 1$:*

1. *Prices satisfy the asset pricing equation* (3)
2. *The dividend evolves by equation* (4)
3. *The information set evolves by equation* (5)

Demonstrating equivalence between nescient and non-nescient equilibria in the finite and infinite history models is most straightforward when considering AR(1) processes for the initial signal s_0:

$$s_t = \phi s_{t+1} + \varepsilon_{t+1} \tag{21}$$

with $|\phi| < 1$. Even though this is a restrictive assumption on the form of possible signals, it is sufficiently flexible to produce a large set of equilibria.[8] Theorem 7 formalizes this: any equilibrium associated with a first-order Blaschke information basis (i.e. the examples studied in Section 3.2) can be rationalized as a nescient equilibrium of the finite history model, by assuming a suitable AR(1) process for the initial signal.

[8] In principle, further equilibria are possible if the signal process is allowed to be more general than an AR(1) process, but this generality needlessly complicates the proof of Theorem 7.

Theorem 7 *If $w_t = W(L)\varepsilon_t$ is the information basis of an infinite history asset pricing model with equilibrium price process p_t, and $W(L)$ is a first-order Blaschke product, then p_t is also the nescient equilibrium price process of some finite history asset pricing model for $t \geq 1$.*

Proof: Appendix A.7

Any infinite history asset pricing model has a continuum of equilibria associated with first-order Blaschke information bases. These equilibria are non-nescient. Theorem 7 says that each of these non-nescient equilibria are exactly equivalent for $t \geq 1$ to the nescient equilibrium of the finite history model, given some initial condition. The strategy for the proof is to first show that the Wold innovations to the AR(1) signal process (21) are a first-order Blaschke information basis. Then I show that observing the current shock and prior period's signal reveals the current signal. Finally, I prove that the nescient price conditional on observing the signal process is exactly the price of a non-nescient equilibrium with the same information basis.

4.2 Nescience Discussion

Axiomatically choosing the nescient equilibrium is not a feasible way to resolve the multiplicity. The multiplicity of equilibria calculated in Section 3.3's examples can all be recast as nescient equilibria of a finite history model. Nescience is not enough to select unique equilibria.

What if this argument is unconvincing and all non-nescient equilibria are discarded anyway? This would just transform the problem of multiplicity into a problem of dependence on initial conditions. The news in the initial information set self-perpetuates; its influence never dissipates. When solving the neoclassical growth model or stationary DSGE models, the initial condition is rarely a concern; while initial conditions matter for equilibrium, they matter less and less as time passes. This is not so for the informational initial condition. The initial information set determines the long run autocovariance function of the economy.

What is the benefit of considering this issue as multiplicity versus initial condition dependence? The infinite history model is more tractable than the finite history model. Finding multiple equilibria is as simple as conjecturing non-causal Blaschke products and testing if they form valid information bases. Whereas in the finite history model, it is more challenging to construct the appropriate initial condition. This was straightforward in the proof of Theorem 7, where the possible set of information bases was restricted to first-order Blaschke products, which are consistent with AR(1) processes for the initial signal. Higher order information bases require constructing more complicated initial signals when the history is finite. So the infinite history model is convenient because it requires no further work to characterize an equilibrium beyond finding the information basis.

The infinite history was not a necessary assumption to generate the multiple equilibria. Is the infinite future? For example, the recursive news signal s_t defined by equation (21) is a discounted sum of all future shocks. But this assumption can be relaxed too. Instead of telescoping forever into the future, equation (21) can recursively determine the signal up to some final stochastic s_T. Nor is it problematic that the information basis is the linear combination of all future shocks; Lemma 1 shows that an information basis can be rewritten in terms of current and past news signals s_t.

Finally, the nescient equilibrium need not be a special equilibrium that justifies axiomatic selection. In Section 3.3's example, the nescient equilibrium was special because it was a lower bound on news. But in some models, there may exist a continuum of equilibria with even less information about the future than the nescient equilibrium. For example, if agents are limited to observing dividends x_t and prices p_t, these series need not causally reveal the fundamental shocks ε_t in all equilibria. Indeed, this is one of the mechanisms by which fundamentally causal models can produce equilibria that feature apparent non-causality. The next section explores such possibilities.

5 Mechanisms for Non-causality

How restrictive is the assumption of causality? Does it make sense for current choices to depend on future shocks? Yes, if future economic shocks are not truly unknowable; such shocks are *nonfundamental.*[9] This section provides several plausible examples of white noise shocks that can be predicted by other strictly causal time series. In these examples, the asset price is such a time series, and the predictability arises endogenously as self-fulfilling news.

[9]Nonfundamentalness is a well known challenge to recovering economic shocks from reduced form analysis (Hansen and Sargent, 1991). Early examples of models featuring nonfundamentalness include Hansen and Sargent (1980), Futia (1981), Quah (1990), and Lippi and Reichlin (1993).

5.1 Non-invertible Shocks

In this section, I present a full information asset pricing model which features multiple *causal* equilibria. The key assumption is that the fundamental shocks ε_t are generated by some underlying unknowable *superfundamental* shock φ_t, such that equilibrium prices may still be causal despite containing news about future ε_t's.[10] And Theorem 8 proves that any non-causal equilibrium studied in Section 3.3 can be represented as a causal equilibrium in terms of some superfundamental.

Definition 7 *A white noise shock process ε_t is said to be determined by a* ***superfundamental*** *shock process φ_t if*

1. *φ_t is a shock process, i.e. φ_t is white noise with finite variance*
2. *ε_t is causally spanned by φ_t, i.e. there exists a causal polynomial $Z(L)$ such that*

$$\varepsilon_t = Z(L)\varphi_t$$

3. *The polynomial $Z(L)$ is not causally invertible.*

What is a superfundamental? Any underlying process that causally determines future news, but cannot be recovered from contemporaneous data. For example, agents may anticipate future productivity improvements, but those future "shocks" may be determined by current changes in technological R&D. The R&D would be the superfundamental process that cannot be recovered from observing contemporaneous productivity.

Crucially, any non-causal equilibrium is observationally equivalent to a causal equilibrium where the news is determined by some superfundamental shock process:

Theorem 8 *If $p_t = P(L)w_t$ is a stationary equilibrium process with information basis $w_t = W(L)\varepsilon_t$, and the price contains some news so that $p_t \neq p_t^N$, then the p_t process*

[10] Using the VAR literature's language, the shocks ε_t may be considered nonfundamental, and φ_t the true fundamental. However, I avoid this language because most studies of nonfundamentalness assume that the fundamental shocks enter the information sets of agents in the model (Beaudry and Portier, 2014). Therefore, I continue to label ε_t the "fundamental" shock, because the "superfundamental" φ_t is never explicitly revealed to agents. This mechanism is closely related to the "confounding dynamics" studied by Rondina and Walker (2021). In their business cycle model, information frictions prevent agents from exactly learning the causal shock process, even though the model features as many shocks as signals.

is one of multiple stationary equilibria causally determined by some superfundamental shock process φ_t.

Proof: Appendix A.8

The intuition of the proof is that given any information basis $w_t = W(L)\varepsilon_t$, it is possible to causally write the shock in terms of some superfundamental by choosing $Z(L) = W^*(L)$.

To illustrate, consider the following example which features multiple equilibria that are causally determined by the superfundamental. The process for the observable fundamental shock is given by a second-order Blaschke product:

$$Z(L) = B_1(L)B_2(L)$$

where $B_1(L)$ and $B_2(L)$ are *causal* Blaschke factors:

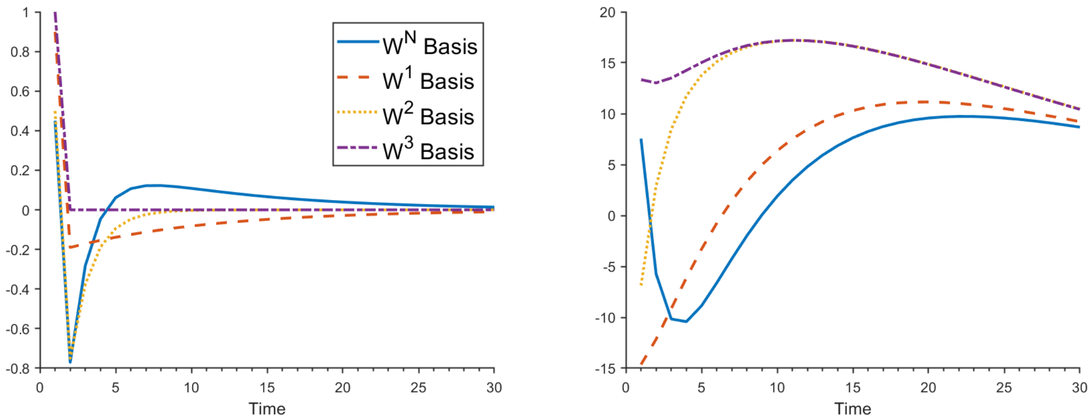
$$B_1(L) = (\theta_1 - L)(1 - \theta_1 L)^{-1} \qquad B_2(L) = (\theta_2 - L)(1 - \theta_2 L)^{-1}$$

The adjoints of any combination of these causal Blaschke factors form valid information bases. The four cases are:

$$W^N(L) = I \qquad W^1(L) = B_1^*(L) \qquad W^2(L) = B_2^*(L) \qquad W^3(L) = B_1^*(L)B_2^*(L)$$

The first valid cast is the nescient equilibrium: $w_t^N = W^N(L)\varepsilon_t = \varepsilon_t$. The other three are combinations of Blaschke adjoints. These information bases effectively "root-flip" one or both of the Blaschke factors that make up $Z(L)$. Higher order Blaschke products would admit even more causal equilibria: one for every combination of Blaschke factors.

Figure 2 presents a numerical example, plotting the four causal equilibria. The dividend x_t is AR(1) and the parameter values are unchanged from Section 3.3. The roots for the Blaschke factors are $\theta_1 = 0.5$ and $\theta_2 = 0.9$. In both panels, the IRFs are with respect to a unit superfundamental shock. Panel (a) gives the four information bases, each of which is causal. The nescient equilibrium reveals the least about the superfundamental, so W^N is the convolution of W^1 and W^2 and has the smallest weight on the contemporaneous shock. W^3 fully reveals the superfundamental, so it only has weight on the contemporaneous shock. Panel (b) plots the IRFs of the prices in terms of the superfundamental. The nescient price has exactly the same autocovariance function as in Figure 1, despite looking dissimilar. The other prices all contain news about future dividends, so their IRFs have larger sums of squares than the nescient



(a) Information IRF: Superfundamental Basis (b) Price IRF: Superfundamental Basis

Figure 2: Impulse Response Functions in the AR(1) Example

IRF, implying higher price volatilities. The highest variance price is for the W^3 basis, which reveals the most information about the superfundamental.

This example of self-fulfilling causal news demonstrates that any non-causal equilibrium can be recast as a causal equilibrium in terms of some superfundamental. It is one way that an observed white noise shock can be determined by past processes. This may result in multiplicity, because those past processes can be revealed as self-fulfilling news by equilibrium prices.

5.2 Non-causality Discussion

When exogenous variables are non-invertible, fundamentally causal models can appear to be non-causal, and feature multiple equilibria. Moreover, this non-invertibility can appear in many more ways than a simple scalar time series with explosive roots. For example, if dividends are determined by multiple superfundamental shocks, then the dividend process is non-invertible and the asset pricing model has multiple equilibria (Appendix B.1). Or, if agents receive some exogenous news about future shocks, and the news signal has a recursive structure, then again the model features multiple equilibria (Appendix B.2).

The non-invertible example also suggests an alternative interpretation to the main finding. Instead of accepting non-causality and concluding that FIRE models generally feature multiple equilibria, economists can draw a weaker conclusion: if agents only have *common information*, where they observe prices and dividends, but not the fundamental shocks driving the dividends, then if dividends are determined by some non-invertible process, the asset market has multiple equilibria.

Non-invertibility examples demonstrate that relaxing the restrictive causality assumption is realistic. Theorem 1 ensures that when this assumption is relaxed, the general asset pricing model has multiple solutions. Does this conclusion about multiplicity carry over to more sophisticated macroeconomic settings? The next section answers this question.

6 Macroeconomic Models

This section considers DSGE models in general, which can feature multiplicity. I show how each equilibrium is characterized by an information basis, as in Section 3.2. As an example, I study the canonical RBC model, for which I calculate many full information equilibria.

6.1 A General DSGE Model

Consider a general linear dynamic stochastic macroeconomic model of the following form. Y_t is a $n \times 1$ vector of endogenous variables, and X_t is a vector of exogenous variables. The equilibrium conditions of the model are represented as a single matrix equation:

$$0 = E[B_{X0}X_t + B_{X1}X_{t+1} + B_{Y0}Y_t + B_{Y1}Y_{t+1}|\Omega_t] \tag{22}$$

with some implicit constraints that any state variables in Y_t are predetermined at time t. The exogenous variables are causally determined by fundamental white noise shocks ε_t:

$$X_t = X(L)\varepsilon_t \tag{23}$$

Agents have full information, so their information set Ω_t evolves by:

$$\Omega_t = \{\Omega_{t-1}, \varepsilon_t, Y_t\} \tag{24}$$

These three equations characterize equilibrium:

Definition 8 *A* ***full information equilibrium of the macroeconomic model*** *is a sequence of endogenous vectors Y_t, exogenous vectors X_t, and information sets Ω_t, given a sequence of shocks ε_t such that for all t:*

1. *Endogenous vectors satisfy the matrix equation* (22)

2. *The exogenous vector evolves by equation* (23)

3. *The information set evolves by equation* (24)

Standard solution methods solve this general DSGE model by conjecturing that the endogenous variables Y_t are a function of current and past variables. Expressed in terms of fundamental shocks, this conjecture implies that $Y_t = Y(L)\varepsilon_t$, with the lag operator polynomial $Y(L)$ causal. Standard methods solve for the *nescient* equilibrium.

But conjecturing that choices are a function of causal variables assumes away other possible equilibria! People make choices as a function of their information, which includes causal shocks, but potentially news as well. In the full information model, Y_t enters the information set, so if Y_t contains news about future shocks, agents will make choices based on that news, which can allow the news to appear in Y_t. As in the asset-pricing model, news begets news, because macroeconomic variables are forward-looking even when there are no asset prices.

Theorem 9 formalizes the existence of multiple equilibria in the general macroeconomic model. Let Y_t^N and Y_t^{PF} denote the equilibrium time series corresponding to nescient information Ω_t^N and perfect foresight Ω_t^{PF}, respectively (equations (8) and (10)). The proof strategy follows exactly the approach used for Theorem 1.

Theorem 9 *If $Y_t^N \neq Y_t^{PF}$ for some t, then there exist multiple stationary full information equilibria of the general DSGE model*

Proof: Appendix A.9

Stationary equilibria are characterized by an information basis, as in Section 3.2. For the macroeconomic model, the information basis W_t is defined:

Definition 9 *A process W_t is an **information basis** for an equilibrium macroeconomic process X_t if:*

1. *ε_t is causally spanned by W_t, i.e. there exists a lag operator polynomial $U(L)$ such that $\varepsilon_t = \sum_{j=0}^{\infty} U_j L^j W_t$*

2. *W_t is the process of innovations for the Wold representation of $X(L)$, i.e. for $X_t = X(L)W_t$, $X(L)$ is both causal and causally invertible, while W_t is unit white noise so that*

$$cov(W_t, W_{t-j}) = \begin{cases} I & j = 0 \\ 0 & j \neq 0 \end{cases}$$

The main difference between this definition and that of the asset pricing model is that the information basis for a macroeconomic model may be vector-valued.[11]

In a stationary equilibrium, endogenous variables are lag operator polynomials of the information basis:

$$Y_t = Y(L)W_t$$

with $Y(L)$ causal. If $W_t = \varepsilon_t$, then Y_t is the nescient equilibrium Y_t^N, which is found by standard methods. Otherwise Adams (2021a) proves how to linearly solve for the equilibrium polynomial $Y(L)$ given an information basis W_t. If the Blanchard-Kahn conditions are satisfied, then the equilibrium process is

$$Y(L)W_t = -\Theta(L)\left[\Xi(L)\left[B_{X1}^{-1}\left(B_{Z1}L^{-1} + B_{Z0}\right)X(L)W^*(L)\right]_+\right]_+ W_t$$

where $\Theta(L)$ and $\Xi(L)$ are known polynomials that depend on the matrix coefficients in equation 22. Crucially, this solution holds for any arbitrary information basis W_t, so Y_t will be an equilibrium process so long as W_t satisfies all of the properties of Definition 9. As the next example demonstrates, there may be many such bases.

6.2 Real Business Cycle Example

In this example, I study multiplicity in a simple RBC model resembling that of Kydland and Prescott (1982).

The linearized model features five equilibrium conditions and as many endogenous variables.The production function gives output y_t in terms of capital k_t, labor l_t, and exogenous productivity a_t:

$$y_t = a_t + \alpha k_t + (1-\alpha)l_t$$

The resource constraint requires that all output is used for consumption c_t and investment i_t:

$$\overline{Y}y_t = \overline{C}c_t + \overline{I}i_t$$

[11]This complicates some findings; in particular I have no general analog to Theorem 4, because the proof relied on lag operator polynomials commuting, which need only be true when they are scalar-valued. Even in the multivariate case, the Wold decomposition still applies, and the noncausal basis polynomial $W(L)$ represents some analytic multivariate inner function that is a product of noncommuting multivariate Blaschkes products and singular inner functions (Jury, Martin, and Shamovich, 2021).

where $\overline{Y}$, $\overline{C}$, and $\overline{I}$ are the steady-state values of output, consumption, and investment respectively. Capital is a state variable, with law of motion

$$k_{t+1} = \delta i_t + (1-\delta)k_t$$

where δ is the depreciation factor. The forward-looking Euler equation for a consumer with log utility is

$$0 = \beta E[c_t - c_{t+1} + \overline{R}(y_{t+1} - k_{t+1})|\Omega_t]$$

where $\overline{R}$ is the steady state gross return on capital. The first order condition for labor supply is

$$y_t = c_t + \frac{1+\eta}{\eta} l_t$$

where η is the Frisch elasticity. These five equations make up the matrix equation (22). Finally, I assume that productivity is AR(1) in terms of the fundamental shock ε_t:

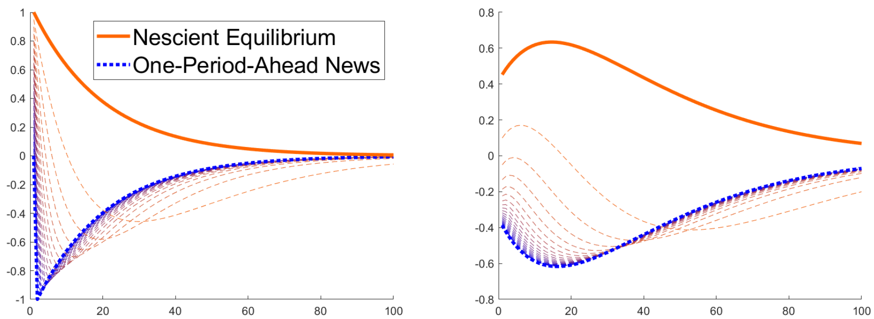
$$a_t = \rho a_{t-1} + \varepsilon_t$$

Parameter	Interpretation	Value
β	Discount factor	0.99
η	Frisch labor supply elasticity	0.5
α	Capital share	0.33
δ	Depreciation factor	0.02
ρ	Productivity autocorrelation	0.95

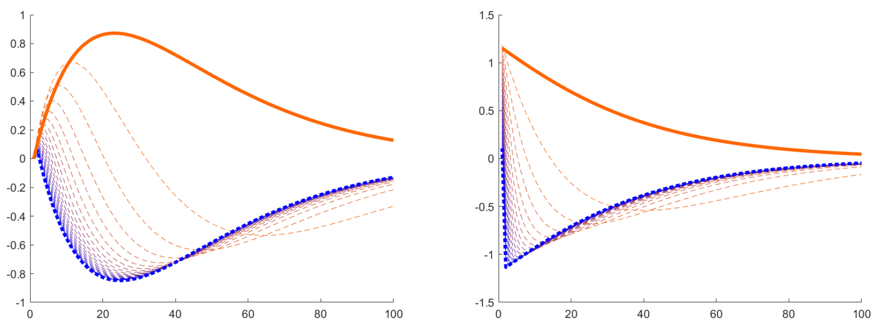
Table 1: Quarterly Calibration

There is only one shock, so the information basis $W_t = W(L)\varepsilon_t$ is scalar-valued in this example. Many possible information bases exist; I calculate equilibria of information bases that are first-order Blaschke factors (equation (16)) selecting 20 values of ϕ, equally spaced on $[0, 1]$. The model's calibrated parameters are otherwise standard, listed in Table 1.

Figure 3 presents the causal IRFs for how four macroeconomic quantities respond to a unit innovation in the information w_t. Panel (a) plots the IRFs for productivity, which become negative as $\phi \to 0$, because one-period-ahead news is the negative of the future shock, i.e. $w_t = -\varepsilon_{t+1}$. Panel (b) plots consumption; as usual, in the one-period-ahead news case, consumption jumps by less on impact because agents



(a) Productivity (b) Consumption

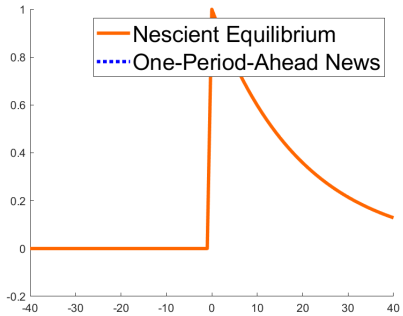


(c) Capital (d) Output

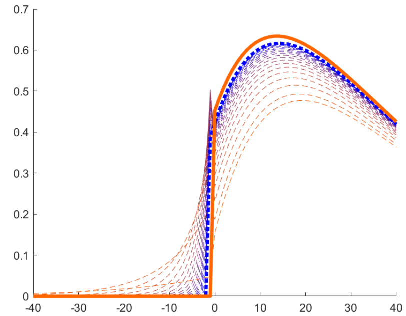
Figure 3: Multiple Equilibria in the RBC Model: IRFs expressed in the Information Bases

are able to improve their consumption smoothing with the addition of news about the future. Panel (c) plots capital, which by assumption cannot jump in response to contemporaneous information. After news is realized, it rises for one period as agents consume out of their capital stock in order to afford their smoothed consumption in advance of future productivity improvements. Panel (d) plots output, which follows the paths of productivity and capital, plus a small amplification that is due to the labor supply response.

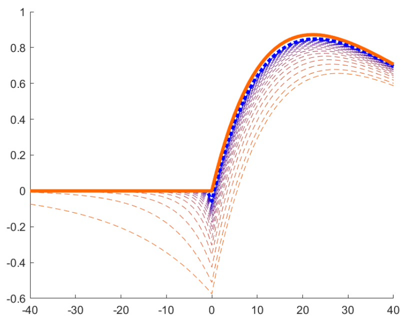
Figure 4 the plots the non-causal IRFs for the same quantities in terms of the fundamental shocks. Panel (a) demonstrates that for any information basis, the shock has exactly the same effect on productivity. Consumption in Panel (b) is forward-looking, so it can respond to news and thus non-causal shocks. Unlike the smooth IRFs in Figure 3, there are jumps in the consumption responses to fundamental shocks. This is because agents cannot distinguish between different future shocks for intermediate



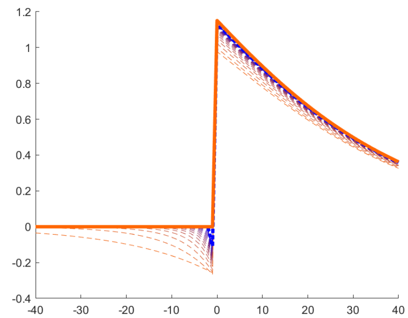
(a) Productivity



(b) Consumption



(c) Capital

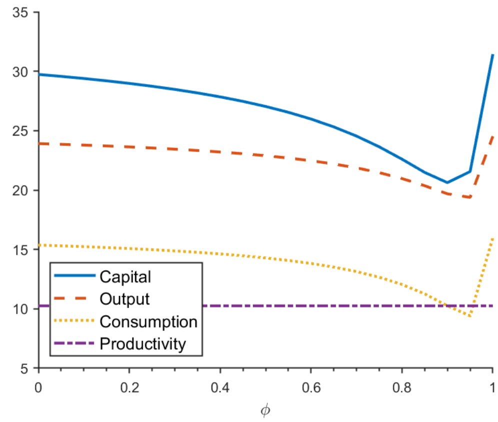


(d) Output

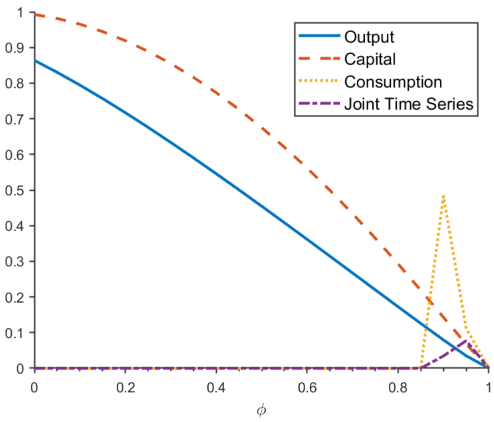
Figure 4: Multiple Equilibria in the RBC Model: IRFs expressed in the Shock Basis

values of ϕ; they can only smooth consumption with respect to their information basis. Capital in Panel (c) now can have a nonzero response to contemporaneous shocks, because they are wrapped up in the news that affected investment in the past. Thus even though the contribution of productivity to the output IRFs in Panel (d) is entirely causal, output is non-causal because capital and labor depend on forward-looking choices.

In the asset pricing model, the additional equilibria increased price volatility relative to the nescient equilibrium; this is not the case for the business cycle model. Figure 5 Panel (a) plots the variance of different time series with respect to the root ϕ that parameterizes the information basis. Productivity cannot be affected by the information basis. For all other time series, the highest volatility is in the nescient equilibrium ($\phi = 1$). As is standard, the variance of consumption and other time series is lower when agents receive one-period-ahead news ($\phi = 0$) because they are able to improve their consumption smoothing. However, the intermediate cases give even



(a) Volatility



(b) Noninvertibility Measure

Figure 5: Sensitivity to the Basis Parameter in the RBC Model

greater attenuation, with the lowest variances for ϕ close to the upper bound. This contrasts with the result in the asset pricing model, where the highest variance was a for $\phi = \beta$. This is because in both cases, a large ϕ implies that the basis contains much information about shocks far into the future. The news increased volatility in the asset pricing model by improving forecasting of future dividends, but this decreases volatility in the RBC model by improving consumption smoothing.

Many equilibria exist for the full information RBC model! However unlike the asset pricing model, not all of the considered information bases were valid. This is because when a time series is a function of a basis, that time series may not be invertible to recover the basis. For example, consumption is given by $c_t = C(L)W_t$, but $C(L)$ may not be causally invertible, in which case W_t would not correspond to the white noise of the Wold representation of c_t, and would not be a valid information basis. To determine whether W_t is a valid basis for the endogenous time series $Y_t = Y(L)W_t$, I check if $Y(L)$ is causally invertible, which is the case for $\phi \leq 0.85$ and $\phi = 1$.

For non-invertible processes, it is possible to calculate how far the process is from being invertible by taking the Wold decomposition of $Y(L) = Z(L)U(L)$, where $Z(L)$ is causally invertible and $U(L)$ is unit white noise. $U(L)$ is causal and when $Z(L)$ is causally invertible, the initial coefficient U_0 is one while all other terms are zero. For each ϕ, I calculate U_0, and let $1 - U_0^2$ denote the "noninvertibility measure". Figure 5 Panel (b) plots this measure for each ϕ as the "Joint Time Series", which is non-zero for $\phi \in (0.85, 1)$. Additionally, I calculate the measure for some individual time series to show how they contribute to invertibility. Consumption appears to be most important; it is invertible whenever the entire vector is. Capital and output are never invertible except for the nescient equilibrium, perhaps because capital is predetermined, and output is driven mostly by capital and productivity, neither of which are invertible when $\phi < 1$.

7 Conclusions and Possible Solutions

Dynamic stochastic FIRE models can feature many equilibria. What should macroeconomists do about it?

One possibility is equilibrium selection. Perhaps there exist defensible axioms that will select a single equilibrium from the continuum of possibilities. It is not clear that the nescient equilibrium - which most of the literature currently selects - is the correct choice given the empirical evidence for the presence of news (Beaudry and Portier, 2014). Additionally, the causal news example in Section 5.1 demonstrates that the nescient equilibrium is not even special as an extreme or limiting case, or by axiomatically discarding non-causal equilibria, because there is also multiplicity as information about past shocks gets removed.

Given the empirical literature, it is tempting to use data to pin down a model; each equilibrium is uniquely (up to scale) associated with an information basis, which can be estimated from the relevant time series. This could be a viable resolution. However, this approach would require first ruling out spontaneous news, as there is no reason yet to rule out having temporary deviations from an information basis, or even to switch entirely from one stationary equilibrium to another.

Another possibility is to embrace the multiplicity. Certainly there are many clear examples of multiple equilibria in macroeconomics, such as bank runs (Diamond and Dybvig, 1983) or currency crises (Obstfeld, 1996). But should we conclude that business cycles in general feature spontaneous shifts from one information process to another? Perhaps this approach could help explain patterns of nonfundamental volatility.

A feasible solution is to relax full information or rational expectations. Incomplete information and bounded rationality can break the feedback from actions to forecasts to actions that creates the multiplicity. The incomplete information literature is particularly suited to address the exact problems raised in this paper, because incomplete information theories jointly determine economic actions and information sets in equilibrium. Sometimes this is done by assuming strictly exogenous processes for information, such as in Woodford (2003), Lorenzoni (2009), or Angeletos and La'O (2013). But it is increasingly common to allow learning from endogenous information, as in Graham and Wright (2010), Melosi (2016), or Adams (2021b). Sometimes such learning does not eliminate the FIRE multiplicity, although it may reduce it to a small finite number of equilibria (Chahrour and Gaballo, 2020) or to locally unique equilibria (Adams, 2021a). In other cases, the multiplicity is erased entirely; Lucas (1972) featured a unique equilibrium.

References

Adams, J. J. (2021a): "Macroeconomic Models with Incomplete Information and Endogenous Signals," Discussion Paper 001004, University of Florida, Department of Economics, Publication Title: Working Papers.

——— (2021b): "Moderating Macroeconomic Bubbles Under Dispersed Information," *University of Florida mimeo.*

Angeletos, G.-M., and J. La'O (2013): "Sentiments," *Econometrica*, 81(2), 739–779.

Beaudry, P., and B. Lucke (2010): "Letting Different Views about Business Cycles Compete," *NBER Macroeconomics Annual*, 24, 413–456, Publisher: The University of Chicago Press.

Beaudry, P., and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 96(4), 1293–1307.

——— (2014): "News-Driven Business Cycles: Insights and Challenges," *Journal of Economic Literature*, 52(4), 993–1074.

Benhabib, J., P. Wang, and Y. Wen (2015): "Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83(2), 549–585.

Blanchard, O. J., and C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–1311.

Chahrour, R., and G. Gaballo (2020): "Learning from House Prices: Amplification and Business Fluctuations," *The Review of Economic Studies.*

Chahrour, R., and K. Jurado (2021): "Recoverability and Expectations-Driven Fluctuations," *The Review of Economic Studies*, p. rdab010.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *The Quarterly Journal of Economics*, 109(1), 241–265.

Diamond, D. W., and P. H. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91(3), 401–419.

Forni, M., L. Gambetti, M. Lippi, and L. Sala (2017): "Noisy News in Business Cycles," *American Economic Journal: Macroeconomics*, 9(4), 122–152.

Futia, C. A. (1981): "Rational expectations in stationary linear models," *Econometrica*, 49(1), 171–192.

Gaballo, G. (2018): "Price Dispersion, Private Uncertainty, and Endogenous Nominal Rigidities," *The Review of Economic Studies*, 85(2), 1070–1110.

Gazzani, A. (2020): "News and noise bubbles in the housing market," *Review of Economic Dynamics*, 36, 46–72.

Graham, L., and S. Wright (2010): "Information, heterogeneity and market incompleteness," *Journal of Monetary Economics*, 57(2), 164–174.

Hansen, L. P., and T. J. Sargent (1980): "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and Control*, 2, 7–46.

Hansen, L. P., and T. J. Sargent (1991): "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent. Westview Press, Boulder.

Jury, M. T., R. T. W. Martin, and E. Shamovich (2021): "Blaschke–singular–outer factorization of free non-commutative functions," *Advances in Mathematics*, 384, 107720.

Kydland, F. E., and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

Lanne, M., and P. Saikkonen (2013): "Noncausal Vector Autoregression," *Econometric Theory*, 29(3), 447–481, Publisher: Cambridge University Press.

Lippi, M., and L. Reichlin (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment," *The American Economic Review*, 83(3), 644–652, Publisher: American Economic Association.

——— (1994): "VAR analysis, nonfundamental representations, blaschke matrices," *Journal of Econometrics*, 63(1), 307–325.

Lorenzoni, G. (2009): "A Theory of Demand Shocks," *American Economic Review*, 99(5), 2050–2084.

Lucas, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of economic theory*, 4(2), 103–124.

Lucas, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46(6), 1429–1445.

Melosi, L. (2016): "Signaling Effects of Monetary Policy," *The Review of Economic Studies.*

Mertens, K., and M. O. Ravn (2010): "Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach," *The Economic Journal*, 120(544), 393–413.

Morris, S., and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *The American Economic Review*, 88(3), 587–597.

Nimark, K. (2017): "Dynamic Higher Order Expectations," *Working Paper.*

Obstfeld, M. (1996): "Models of currency crises with self-fulfilling features," *European Economic Review*, 40(3), 1037–1047.

Prescott, E. C., and R. Mehra (1980): "Recursive Competitive Equilibrium: The Case of Homogeneous Households," *Econometrica*, 48(6), 1365–1379.

Quah, D. (1990): "Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption," *Journal of Political Economy*, 98(3), 449–475, Publisher: The University of Chicago Press.

Rondina, G., and T. B. Walker (2021): "Confounding Dynamics," *Journal of Economic Theory*, 196, 105251.

Singleton, K. J. (1987): "Asset prices in a time-series model with disparately informed, competitive traders," .

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989): *Recursive methods in economic dynamics.* Harvard University Press.

Woodford, M. (2003): "The Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford.

A Proofs

A.1 Theorem 1

Proof of Theorem 1. It suffices to find any valid alternative to the nescient equilibrium.

Let $\tilde{\Omega}_t \supset \Omega_t^N$ denote any information set containing the nescient information set that induces a stationary forecast satisfying

$$E\left[\sum_{j=0}^{\infty}\beta^j x_{t+j}|\Omega_t^N\right] \neq E\left[\sum_{j=0}^{\infty}\beta^j x_{t+j}|\tilde{\Omega}_t\right] \tag{25}$$

with associated stationary price $\tilde{p}_t = E[\sum_{j=0}^{\infty}\beta^j x_{t+j}|\tilde{\Omega}_t]$.

The implied information set is

$$\Omega_t = \{\Omega_{t-1}^N, \varepsilon_t, \tilde{p}_t\}$$

$\tilde{p}_t$ is the best forecasts of $\sum_{j=0}^{\infty}\beta^j x_{t+j}$ conditional on $\tilde{\Omega}_t$, so it must also be the best forecast conditional on Ω_t, which adds no additional information. Therefore $p_t = \tilde{p}_t$. The assumption that $\tilde{p}_t \neq p_t^N$ implies that $\Omega_t \neq \Omega_t^N$ and $p_t \neq p_t^N$. The time series Ω_t and p_t make up a valid equilibrium, but so do Ω_t^N and p_t^N.

Finally, it remains to show that such a $\tilde{\Omega}_t$ exists. One candidate is the perfect foresight information set Ω_t^{PF}. By assumption $p_t^N \neq p_t^{PF}$, so Ω_t^{PF} must satisfy condition (25). ■

A.2 Theorem 2

Proof of Theorem 2. First I prove that $w_t = W(L)\varepsilon_t$ can be written as a polynomial in current and future ε_t alone, i.e. $W(L)$ is anti-causal so that coefficients on lag operators satisfy $W_j = 0$ for all $j \geq 1$.

Definition 5 requires that w_t causally spans the shock process by $\varepsilon_t = U(L)w_t$ for some $U(L)$, so the covariance must also be given by

$$cov(w_t, L^j\varepsilon_t) = cov(w_t, L^jU(L)w_t)$$

$$= cov(w_t, \sum_{k=-\infty}^{\infty}U_kL^{k+j}w_t) = cov(w_t, U_{-j}w_t)$$

$$= U_{-j} = 0 \quad \forall j > 0$$

where the last equality holds because $U(L)$ must be causal. Thus $W_j = 0$ for all $j > 0$.

To see that $W(L)$ is invertible by its adjoint, consider the polynomial given by

$$V(L) = W^*(L)W(L)$$

Each coefficient is

$$\begin{aligned} V_j &= \sum_{k=-\infty}^{\infty} W^*_{-k} W_{k+j} = \sum_{k=-\infty}^{\infty} W_k W_{k+j} \\ &= cov(W(L)\varepsilon_t, L^j W(L)\varepsilon_t) \\ &= cov(w_t, w_{t-j}) = \begin{cases} 0 & j \neq 0 \\ 1 & j = 0 \end{cases} \end{aligned}$$

where the last equality follows from the white noise property of w_t. This implies that $V(L)$ is the identity. ■

A.3 Theorem 3

Proof of Theorem 3. The pricing equation is

$$p_t = x_t + \beta E[p_{t+1}|\Omega_t]$$

which when written with lag operator polynomials becomes

$$P(L)w_t = X(L)\varepsilon_t + \beta E[P(L)w_{t+1}|\Omega_t]$$

The information set Ω_t is the history of the information basis, i.e. $\Omega_t = \{w_{t-j}\}_{j=0}^{\infty}$. Therefore $E[w_{t+1}|\Omega_t] = 0$ which implies the forecast is given by

$$E[P(L)w_{t+1}|\Omega_t] = E[\sum_{j=0}^{\infty} P_j L^j w_{t+1}|\Omega_t]$$

$$= \sum_{j=1}^{\infty} P_j L^j w_{t+1} = [P(L)L^{-1}]_+ w_t$$

using the annihilation operator. Plugging this forecast back into the pricing equation:

$$P(L)w_t = X(L)\varepsilon_t + \beta[P(L)L^{-1}]_+ w_t$$

Use the relationship $\varepsilon_t = W^*(L)w_t$:

$$P(L)w_t = X(L)W^*(L)w_t + \beta[P(L)L^{-1}]_+ w_t$$

Collecting coefficients on w_t gives a relationship for the polynomials:

$$P(L) = X(L)W^*(L) + \beta[P(L)L^{-1}]_+$$

Substituting recursively yields

$$P(L) = X(L)W^*(L) + [\beta L^{-1}X(L)W^*(L) + \beta^2 L^{-2}X(L)W^*(L) + ...]_+$$

$$= [(I - \beta L^{-1})^{-1}X(L)W^*(L)]_+$$

■

A.4 Theorem 4

Proof of Theorem 4. When $\phi = 1$, $(1 - \phi L^{-1})$ is not invertible, so the limit must be taken. The limiting information basis is

$$\lim_{\phi\to 1} w_t = \lim_{\phi\to 1}(\phi - L^{-1})(1 - \phi L^{-1})^{-1}\varepsilon_t$$

$$= \lim_{\phi\to 1}(\phi - L^{-1})(1 + \phi L^{-1} + \phi^2 L^{-2} + ...)\varepsilon_t$$

$$= \lim_{\phi\to 1}(\phi + (\phi^2 - 1)L^{-1} + (\phi^2 - 1)\phi L^{-2} + (\phi^2 - 1)\phi^2 L^{-3} + ...)\varepsilon_t$$

$$= \varepsilon_t$$

and the information set becomes

$$[\phi = 1]: \qquad \Omega_t = \{w_{t-j}\}_{j=0}^{\infty} = \{\varepsilon_{t-j}\}_{j=0}^{\infty}$$

which is the nescient information set Ω_t^N.

When $\phi = 0$, the information basis becomes

$$[\phi = 0]: \qquad w_t = -L^{-1}\varepsilon_t = \varepsilon_{t+1}$$

and the information set becomes

$$[\phi = 0]: \qquad \Omega_t = \{w_{t-j}\}_{j=0}^{\infty} = \{\varepsilon_{t+1-j}\}_{j=0}^{\infty}$$

which says that agents at time t receive perfect news about the one-period-ahead shock.

When $\phi = \beta$, the information basis becomes

$$[\phi = \beta]: \qquad w_t = (\beta - L^{-1})(1 - \beta L^{-1})^{-1}\varepsilon_t$$

Plug this basis into the solution (Theorem 3) with $U = (1 - \beta L^{-1})(\beta - L^{-1})^{-1}$ to get the price process $P_\beta(L)$:

$$P_\beta(L) = [(1 - \beta L^{-1})^{-1} X(L)(1 - \beta L^{-1})(\beta - L^{-1})^{-1}]_+$$
$$= [X(L)(\beta - L^{-1})^{-1}]_+ = X(L)(\beta - L^{-1})^{-1}$$

because $(\beta - L^{-1})^{-1} = (\beta L - 1)^{-1} L$ which is strictly causal, so the annihilation operator can be dropped. Multiply by $(1 - \beta L^{-1})^{-1}(1 - \beta L^{-1})$ to get

$$= (1 - \beta L^{-1})^{-1} X(L)(1 - \beta L^{-1})(\beta - L^{-1})^{-1} = (1 - \beta L^{-1})^{-1} X(L) U(L)$$

then multiply both sides by $U^{-1}(L)\varepsilon_t$ to get

$$P_\beta(L) U^{-1}(L)\varepsilon_t = (1 - \beta L^{-1})^{-1} X(L)\varepsilon_t$$

The left-hand side is simply $P_\beta(L) w_t$. Expand the right-hand side to recover the perfect foresight price (equation (11)):

$$P_\beta(L) w_t = (1 - \beta L^{-1})^{-1} x_t = \sum_{j=0}^{\infty} \beta^j x_{t+j} = p_t^{PF}$$

■

A.5 Theorem 5

Proof of Theorem 5. The process for the information basis w_t is invertible by its adjoint $W^*(L)$, so the fundamental shocks can be written in terms of the basis by

$$\varepsilon_t = W^*(L) w_t = \sum_{j=0}^{\infty} W_{-j} w_{t-j}$$

which implies that the contemporaneous innovation w_t can be expressed in terms of the contemporaneous shock and past innovations by

$$w_t = \varepsilon_t - \sum_{j=1}^{\infty} W_{-j} w_{t-j} \tag{26}$$

Iterating equation (17) backwards, the information set Ω_t is given by

$$\Omega_t = \{\varepsilon_{t-j}\}_{j=0}^{\infty} \cup \{p_{t-j}^{PO}\}_{j=1}^{\infty}$$

But both $\varepsilon_{t-j} = W^*(L)w_{t-j}$ and $p^{PO}_{t-j} = P(L)^{-1}w_{t-j}$ are causally spanned by the information basis, so the information set is equivalently given by

$$\Omega_t = \varepsilon_t \cup \{w_{t-j}\}_{j=1}^{\infty}$$

so equation (26) implies that $w_t \in \Omega_t$. Therefore the information set Ω_t is equivalent to the information set $\{w_{t-j}\}_{j=0}^{\infty}$ of the perfect-observation model. ■

A.6 Theorem 6

Proof of Theorem 6. Let $\tilde{w}_t = \begin{pmatrix} w_t \\ \nu_t \end{pmatrix}$ denote the two-dimensional information basis with lag operator polynomial $\tilde{W}(L) = \begin{pmatrix} W(L) & 0 \\ 0 & 1 \end{pmatrix}$ determined by

$$\tilde{w}_t = \tilde{W}(L) \begin{pmatrix} \varepsilon_t \\ \nu_t \end{pmatrix}$$

By construction, the information basis $\tilde{w}_t$ causally spans the fundamental shock ε_t, the noise shock ν_t, and the price p_t^{PO}. Therefore it also spans the signal s_t corresponding to the price sequence p_t^{PO} per equation (18), and thus the information set Ω_t per equation (19). Crucially, the entire vector $\tilde{w}_t$ is causally invertible from the information set because $W(L)$ is causally invertible by $W^*(L)$ (Theorem 2):

$$\begin{pmatrix} \varepsilon_t \\ s_t \end{pmatrix} = \begin{pmatrix} \varepsilon_t \\ p_t + \nu_t \end{pmatrix} = \begin{pmatrix} W^*(L) & 0 \\ P(L) & 1 \end{pmatrix} \tilde{w}_t$$

$$\implies \tilde{w}_t = \begin{pmatrix} W^*(L) & 0 \\ P(L) & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_t \\ s_t \end{pmatrix}$$

The equilibrium price given by equation (7) is

$$p_t = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t\right] = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \{\tilde{w}_{t-j}\}_{j=0}^{\infty}\right]$$

$$= E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \{w_{t-j}\}_{j=0}^{\infty} \cup \{\nu_{t-j}\}_{j=0}^{\infty}\right] = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \{w_{t-j}\}_{j=0}^{\infty}\right] = p_t^{PO}$$

where the last step follows from ν_t being orthogonal to w_t and x_t at all lags. ■

A.7 Theorem 7

Lemma 1 reveals how the causal Wold representation of the signal is related to the polynomial $S(L)$ defined by equation (20). I write the signal's Wold representation as $s_t = A^s(L)W^s(L)\varepsilon_t$ with causally invertible $A^s(L)$ and white noise innovation $w_t^s = W^s(L)\varepsilon_t$.

Lemma 1 *If $S(L)$ is a causally invertible rational polynomial, then the recursive news signal s_t has Wold representation $s_t = A^s(L)W^s(L)\varepsilon_t$ with causally invertible $A^s(L)$ given by*

$$A^s(L) = S(L) \tag{27}$$

and unitary $W^s(L)$ given by

$$W^s(L) = L^{-1}S(L)^{-1}S(L^{-1}) \tag{28}$$

Proof of Lemma 1. It is straightforward to show that $s_t = A^s(L)W^s(L)\varepsilon_t$:

$$s_t = S(L^{-1})L^{-1}\varepsilon_t$$

$$= S(L)S(L)^{-1}S(L^{-1})L^{-1}\varepsilon_t = A^s(L)W^s(L)\varepsilon_t$$

and $A^s(L) = S(L)$ is causally invertible by assumption, so it remains to show that $W^s(L)\varepsilon_t$ is white noise.

$S(L)$ is both causal and causally invertible, so $S(L^{-1})$ can be written as

$$S(L^{-1}) = \frac{\prod_{k=1}^{K}(1-\theta_k L^{-1})}{\prod_{j=1}^{J}(1-\rho_j L^{-1})}$$

with J and K finite by assumption, and all roots θ_k and poles ρ_j inside the unit circle.

Therefore $W^s(L)$ can be written as

$$W^s(L) = L^{-1}S(L)^{-1}S(L^{-1}) = L^{-1}\frac{\prod_{j=1}^{J}(1-\rho_j L)}{\prod_{k=1}^{K}(1-\theta_k L)}\frac{\prod_{k=1}^{K}(1-\theta_k L^{-1})}{\prod_{j=1}^{J}(1-\rho_j L^{-1})}$$

Multiply the numerator and denominator of the non-causal polynomial by powers of L:

$$= L^{-1}\frac{\prod_{j=1}^{J}(1-\rho_j L)}{\prod_{k=1}^{K}(1-\theta_k L)}L^{J-K}\frac{\prod_{k=1}^{K}(L-\theta_k)}{\prod_{j=1}^{J}(L-\rho_j)}$$

$$= L^{J-K-1}\frac{\prod_{k=1}^{K}(L-\theta_k)}{\prod_{k=1}^{K}(1-\theta_k L)}\frac{\prod_{j=1}^{J}(1-\rho_j L)}{\prod_{j=1}^{J}(L-\rho_j)}$$

which is a non-causal Blaschke product, so $W^s(L)\varepsilon_t$ is white noise. ■

The most challenging step in proving Theorem 7 is constructing an initial condition that rationalizes a given information basis. Lemma 2 proves how to do so.

Lemma 2 *If the initial information set is given by $\Omega_0 = s_0$ for an AR(1) signal process s_t following*

$$s_t = \phi s_{t+1} + \varepsilon_{t+1}$$

then the nescient equilibrium of the finite history asset pricing model has equilibrium price process $p_t = P(L)\varepsilon_t$ for $t \geq 1$ where $P(L)$ satisfies the pricing equation (13) *with information basis $W(L) = W^S(L)$.*

Proof of Lemma 2. Observing s_{t-1} and ε_t reveals s_t:

$$s_t = \frac{1}{\phi}(s_{t-1} - \varepsilon_t)$$

and observing s_t and s_{t-1} reveals w_t^s by

$$w_t^s = s_t - \phi s_{t-1}$$

because Lemma 1 implies that $(1-\phi L)^{-1}W_t^s$ is the Wold representation for s_t. Therefore, if s_{t-1} and ε_t are in the information set Ω_t, then so are s_t and w_t^s.

s_0 and ε_1 are in Ω_1, thus s_1 and w_1^s are as well. So by induction, Ω_t contains all s_{t-j} for $0 \leq j \leq t$. w_t^s causally spans all s_{t-j} and ε_{t-j} for $j \geq 0$, so w_t^s is a basis for the information set Ω_t.

Agents price using equation (6), which written in terms of shocks is:

$$p_t = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t\right]$$

$$= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \beta^j X_k E\left[\varepsilon_{t+j-k}|\Omega_t\right]$$

The information set can be written $\Omega_t = \{s_0, w_1^S, ...w_t^S\}$. To derive the expectation of future shocks conditional on this information set, first consider the expectation conditional on the infinite history information set $\tilde{\Omega}_t = \{w_{t-j}^S\}_{j=0}^{\infty}$:

$$E[\varepsilon_{t+k}|\tilde{\Omega}_t] = [L^{-k}W^{S*}(L)]_+ w_t^S$$

$$= \begin{cases} \varepsilon_{t+k} & k \leq 0 \\ \sum_{j=0}^{\infty} W^S_{k+j} w^s_{t-j} & k > 0 \end{cases}$$

By Lemma 1, $W^S(L)$ is given by

$$W^S = L^{-1}S(L)^{-1}S(L^{-1}) = L^{-1}\frac{1-\phi L}{1-\phi L^{-1}}$$

$$= -\frac{\phi - L^{-1}}{1-\phi L^{-1}}$$

$W^S(L)$ is first order Blaschke with parameter ϕ, so for $k > 0$

$$[L^{-k}W^{S*}(L)]_+]w^S_t = -[L^{-k}(\frac{\phi - L^{-1}}{1-\phi L^{-1}})^*]_+ w^S_t = -[L^{-k}\frac{\phi - L}{1-\phi L}]_+ w^S_t$$

$$= [L^{-k+1}\frac{1}{1-\phi L} - L^{-k}\frac{\phi}{1-\phi L}]_+ w^S_t = \frac{\phi^{k-1}}{1-\phi L}w^S_t - \frac{\phi^{k+1}}{1-\phi L}w^S_t$$

$$= \phi^{k-1}(1-\phi^2)\frac{1}{1-\phi L}w^S_t = \phi^{k-1}(1-\phi^2)A^S(L)w^S_t$$

$$= \phi^{k-1}(1-\phi^2)s_t$$

The expected value of future shocks conditional on the infinite history information set $\tilde{\Omega}_t$ can be written in terms of the current signal s_t, which is in the finite history information set Ω_t. Therefore, expectations of shocks are the same between the two information sets:

$$E[\varepsilon_{t+k}|\tilde{\Omega}_t] = E[\varepsilon_{t+k}|\Omega_t]$$

and so expectations of future dividends and hence prices are the same between the two information sets:

$$p_t = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j}|\Omega_t\right] = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j}|\tilde{\Omega}_t\right]$$

Theorem 3 says that the pricing polynomial $P(L)$ for the infinite history information set is given by $P(L) = [(I - \beta L^{-1})^{-1}X(L)W^{S*}(L)]_+$, therefore the equilibrium price for the finite history model is

$$p_t = P(L)w^S_t \qquad t \geq 1$$

■

Proof of Theorem 7. For any valid information basis, the infinite history asset pricing model has an equilibrium price process given by

$$p_t = P(L)w_t$$

where $P(L)$ is given by Theorem 3. Let ϕ denote the pole of the non-causal first-order Blaschke factor $W(L)$, and let s_t denote the signal process generated by equation (20) with $S(L^{-1}) = (1 - \phi L^{-1})^{-1}$. If the initial condition of a finite history model is given by $\Omega_0 = s_0$, then Lemma 2 implies that the nescient equilibrium price $\tilde{p}_t$ of the finite history model satisfies

$$\tilde{p}_t = p_t \qquad \forall t \geq 1$$

■

A.8 Theorem 8

Proof of Theorem 8. First I show that there exists a superfundamental process φ_t such that p_t is causal.

$$p_t = P(L)w_t = P(L)W(L)\varepsilon_t$$

Let $\varepsilon_t = Z(L)\varphi_t$ where $Z(L)$ is the causal polynomial to be found:

$$= P(L)W(L)Z(L)\varphi_t$$

Many valid $Z(L)$ exist such that p_t is causal. A simple choice is the adjoin of the information basis. $W^*(L)$ is causal per Definition 5, and inverts $W(L)$ per Theorem 2, so choosing $Z(L) = W^*(L)$ gives:

$$p_t = P(L)W(L)W^*(L)\varphi_t = P(L)\varphi_t$$

By Theorem 3 $P(L)$ is causal, therefore p_t is causally determined by the superfundamental φ_t.

It remains to show that multiple causal equilibria exist for this choice of φ_t. Consider the nescient equilibrium given by $p_t^N = P^N(L)\varepsilon_t$. Substitute for the shock:

$$p_t^N = P^N(L)Z(L)\varphi_t$$

$$= P^N(L)W^*(L)\varphi_t$$

$P^N(L)$ and $W^*(L)$ are both causal so their product is causal. ■

A.9 Theorem 9

Proof of Theorem 9. It suffices to find any valid alternative to the nescient equilibrium.

Rearrange the equilibrium condition (22) as an expression for Y_t:

$$Y_t = E[-B_{Y0}^{-1}B_{Y1}Y_{t+1} - B_{Y0}^{-1}B_{X0}X_t - B_{Y0}^{-1}B_{X1}X_{t+1}|\Omega_t]$$

Substitute for future values Y_{t+1}, Y_{t+2} and so forth to express Y_t in terms of expectations of exogenous vectors:

$$Y_t = -E[\sum_{j=0}^{\infty}(-B_{Y0}^{-1}B_{Y1})^j \left(B_{Y0}^{-1}B_{X0}X_{t+j} + B_{Y0}^{-1}B_{X1}X_{t+1+j}\right)|\Omega_t]$$

Let $\tilde{\Omega}_t \supset \Omega_t^N$ denote any information set containing the nescient information set that induces a stationary forecast satisfying

$$E\left[\sum_{j=0}^{\infty}(-B_{Y0}^{-1}B_{Y1})^j \left(B_{Y0}^{-1}B_{X0}X_{t+j} + B_{Y0}^{-1}B_{X1}X_{t+1+j}\right)|\Omega_t^N\right] \neq E\left[\sum_{j=0}^{\infty}(-B_{Y0}^{-1}B_{Y1})^j \left(B_{Y0}^{-1}B_{X0}X_{t+j} + B_{Y0}^{-1}B_{X1}X_{t+1+j}\right)|\tilde{\Omega}_t\right] \quad (29)$$

with associated stationary vector $\tilde{Y}_t = -E\left[\sum_{j=0}^{\infty}(-B_{Y0}^{-1}B_{Y1})^j \left(B_{Y0}^{-1}B_{X0}X_{t+j} + B_{Y0}^{-1}B_{X1}X_{t+1+j}\right)|\tilde{\Omega}_t\right]$.

The implied information set is

$$\Omega_t = \{\Omega_{t-1}^N, \varepsilon_t, \tilde{Y}_t\}$$

$\tilde{Y}_t$ is the best forecasts of $\sum_{j=0}^{\infty}(-B_{Y0}^{-1}B_{Y1})^j \left(B_{Y0}^{-1}B_{X0}X_{t+j} + B_{Y0}^{-1}B_{X1}X_{t+1+j}\right)$ conditional on $\tilde{\Omega}_t$, so it must also be the best forecast conditional on Ω_t, which adds no additional information. Therefore $Y_t = \tilde{Y}_t$. The assumption that $\tilde{Y}_t \neq Y_t^N$ implies that $\Omega_t \neq \Omega_t^N$ and $Y_t \neq Y_t^N$. The time series Ω_t and Y_t make up a valid equilibrium, but so do Ω_t^N and Y_t^N.

Finally, it remains to show that such a $\tilde{\Omega}_t$ exists. One candidate is the perfect foresight information set Ω_t^{PF}. By assumption $Y_t^N \neq Y_t^{PF}$, so Ω_t^{PF} must satisfy condition (29). ■

B Online Appendix: Additional Mechanisms for Non-Causality

This appendix includes additional examples with apparent non-causality in causal models, and the associated multiplicity of equilibria.

B.1 Shock Complexity

In this example, the dividend process is determined by multiple underlying stochastic processes. Dividend shocks remain white noise, but the shock is not fundamentally unpredictable. The underlying processes are not exogenously revealed to agents, but prices may endogenously contain information about them, manifesting as news about future dividends. Such news is self-fulfilling, allowing for multiple equilibria.

Dividends x_t are now determined by two independent stochastic processes x_t^u and x_t^v:

$$x_t = x_t^u + x_t^v$$

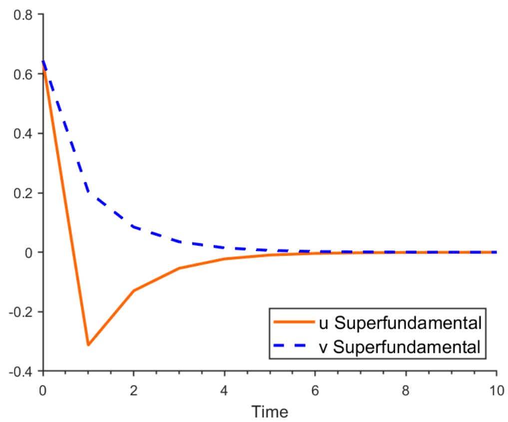
where $x_t^u = X_u(L)u_t$ and $x_t^v = X_v(L)v_t$ are each determined by independent stochastic "superfundamental" shocks u_t and v_t. In the following numerical example, I assume that both processes are AR(1):

$$x_t^u = \rho^u x_{t-1}^u + u_t \qquad x_t^v = \rho^v x_{t-1}^v + v_t \tag{30}$$

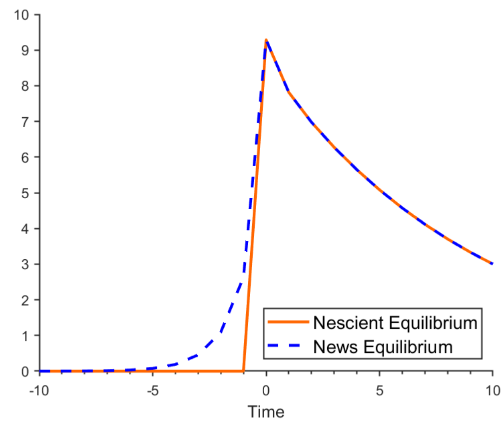
What are these superfundamental shocks? Dividends are determined by a large variety of forces (e.g. demand for different goods in different markets, competitors' actions along different dimensions, behavior of many employees, etc.) that cannot possibly be fully captured by the few data dimensions that traders observe. For this complexity to hide relevant information, there must be more linearly independent superfundamentals than observables: in this simple example, agents observe a single dividend series driven by two independent processes.

The shock ε_t still determines the dividend by $x_t = X(L)\varepsilon_t$, which is now the Wold representation, with ε_t the Wold innovation that is determined by a linear combination of current and past superfundamental shocks. Figure 6 Panel (a) plots this linear combination for the case where $\rho_u = 0.9$ and $\rho_v = 0.1$. The solid orange line is the IRF of ε_t to the superfundamental u_t, while the dashed blue line is the IRF to the superfundamental v_t. The shock ε_t is white noise, but is determined by both current and past values of u_t and v_t because the two processes in equation (30) have different autocorrelations.

Agents in the model observe the dividends x_t and shocks ε_t, but not the underlying superfundamentals u_t and v_t which cannot be inferred by observing x_t alone. If they could observe the superfundamentals, agents could forecast future values of the shock ε_t. Figure 6 Panel (a) makes this clear: ε_t is white noise, but is correlated with past values of u_t and v_t.



(a) ε_t Shock IRF to Each Superfundamental



(b) Price IRF to ε_t Shock

Figure 6: Multiple Equilibria with Shock Complexity

There are at least two causal equilibria to this model. One is the nescient equilibrium, in which prices contain no information beyond current and past values of the shock ε_t. As usual, the nescient price p_t^N is given by equation (9). The second is an equilibrium with self-fulfilling news: prices contain information about future shocks by revealing the current values of the superfundamentals u_t and v_t. This "news equilibrium" price p_t^{ne} now depends independently on each superfundamental, which by equation (6) gives:

$$p_t^{ne} = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j} | \Omega_t\right] = E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j}^u | \{u_{t-j}\}_{j=0}^{\infty}\right] + E\left[\sum_{j=0}^{\infty} \beta^j x_{t+j}^v | \{v_{t-j}\}_{j=0}^{\infty}\right]$$

This price is strictly causal in terms of the superfundamentals, but non-causal when expressed in terms of shocks.

Figure 6 Panel (b) contrasts these two equilibria by plotting the impulse response of prices to ε_t. In the nescient equilibrium, the price depends only on current and past shocks. But in the equilibrium with news, the price is correlated with future shocks, because the price contains information about the superfundamentals not contained in ε_t. The responses of prices to current and past shocks are the same, which is why the IRFs overlap in the causal region. This is because any additional information revealed by prices in the news equilibrium is uncorrelated with information contained in the past history of white noise shocks.

B.2 Exogenous News

In this example, agents receive exogenous news about future shocks, in addition to any endogenous self-fulfilling news. The news process is correlated with future shocks, which are not fundamentally unpredictable, and contains the standard one-period-ahead news as a special case. This example demonstrates two valuable conclusions: FIRE models with exogenous news processes can still feature multiple equilibria, and exogenous news can contain hidden information about future shocks that is revealed in equilibrium.

I introduce exogenous news by modifying the information set evolution equation (5) to:

$$\Omega_t = \{\Omega_{t-1}, \varepsilon_t, p_t, \nu_t\} \tag{31}$$

where ν_t is the exogenous news process, some time series that is correlated with future fundamental shocks. The other equilibrium conditions (3) and (4) are unchanged from the baseline model in Section 3.1.

The presence of exogenous news constrains the set of possible information bases. A valid information basis must now causally span the shocks ε_t, the price process p_t, and the news ν_t. Still, many possible equilibria may exist. To demonstrate, I consider again the AR(1) example, where dividends are given by equation (15). Additionally, I let $\nu_t = \nu(L)\varepsilon_t$ be white noise, with associated lag operator polynomial $\nu(L)$ that is a non-causal first order Blaschke product with parameter ψ. As a result, the news ν_t causally spans the fundamental shocks ε_t (i.e. this news process would be a valid information basis for the example without exogenous news).

This exogenous news process has a recursive structure:

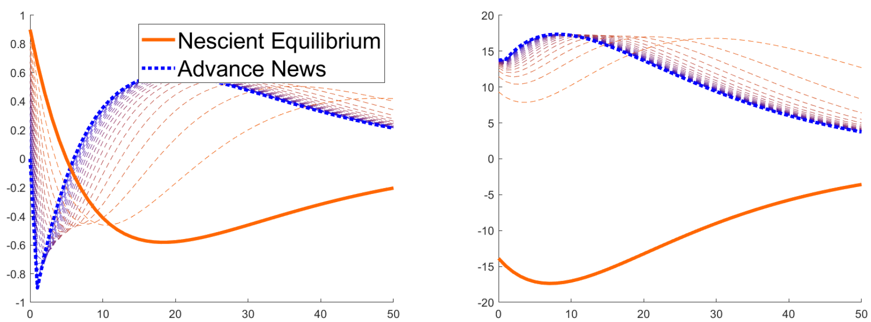
$$\nu_t = \psi\varepsilon_t - \varepsilon_{t+1} + \psi\nu_{t+1}$$

Agents observe the current shock ε_t, so current news ν_t is a signal about the next shock ε_{t+1} and future news ν_{t+1}. These two components cannot be distinguished from one another, except in the $\psi = 0$ special case where news perfectly reveals the next period's shock. When $\psi \in (0, 1)$, the news is correlated with all future shocks.

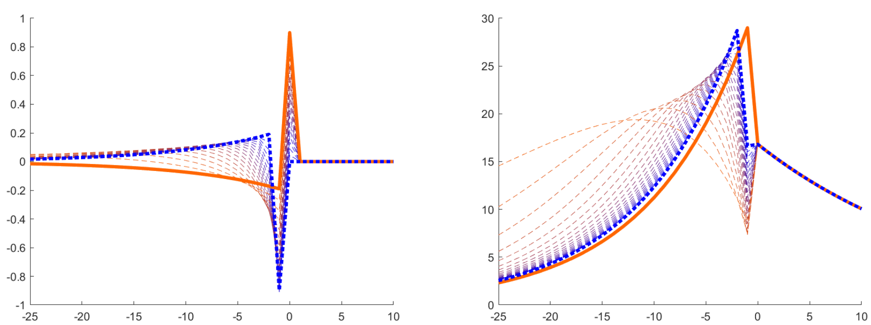
A candidate information basis is given by

$$W(L) = \nu(L)B(L)$$

where $B(L)$ is some non-causal Blaschke product. This information basis is the analog to equation (16) in Section 3.3. It is white noise, causally spanning ν_t and by extension ε_t. If w_t is also the Wold innovation for p_t in equilibrium, then it is a valid information basis (Definition 5).



(a) Dividend IRF: Information Basis (b) Price IRF: Information Basis



(c) Information IRF: Shock Basis (d) Price IRF: Shock Basis

Figure 7: Multiple Equilibria with Exogenous News

I calculate multiple equilibria where $B(L)$ is a first order Blaschke product with parameter ϕ taking twenty values on the unit interval, while $\rho = 0.9$ as in Section 3.3 and $\psi = 0.9$. The equilibria are presented in Figure 7, which is the analog to Figure 1. Two special cases are highlighted. $\phi = 1$ corresponds to the nescient equilibrium, which in this case includes current and past shocks, in addition to current and past news ν_t.

When $\phi = 0$, agents receive news n_{t+1} one period early, which I label "Advance News". This is analogous to the case in Section 15, where $\phi = 0$ implied that agents observed the fundamental shock one period early.

Panel (a) of Figure 7 presents the impulse response functions of dividends to an innovation in the information basis w_t. While the IRFs look dissimilar, they all correspond to exactly the same dividend process in terms of fundamental shocks. Panel (b) gives the IRFs of asset prices to w_t, which as usual is determined by Theorem 3.

Panel (c) of Figure 7 presents the IRFs of the information bases w_t to a fundamental shock ε_t, i.e. exactly the coefficients in the polynomial $W(L)$. The basis corresponding to the nescient equilibrium is now the non-causal Blaschke factor $\nu(L)$, which gives the news process ν_t in terms of the fundamental shocks. Every other information basis is the product of $\nu(L)$ with another non-causal Blaschke factor, creating a second order Blaschke product. As information bases become more complicated, the equilibrium price does as well: Panel (d) presents the IRFs of the asset price to the shock ε_t. The nescient case resembles one of the intermediate cases from Figure 1, while "Advance News" looks similar except shifted one period into the future. The remaining IRFs map out a continuous transformation between the two extremes, which sometimes contain much more news than the exogenous process, and thus place considerable weight on future shocks.

C Online Appendix: Decentralized Nonlinear Models

The linear models considered so far are approximations of nonlinear models. Does the multiplicity found in the linear models also feature in the nonlinear ones? Yes it can. In this section I examine stochastic dynamic programming problems and argue that regularity assumptions which might ordinarily ensure uniqueness fail to do so in decentralized economies.

A general stochastic dynamic programming problem is expressed with a Bellman equation given by

$$V(x, s) = \max_u r(x, s, u) + \beta \int_{s' \in \mathcal{S}} V(g(x, s, u), s') d\lambda(s'|s) \tag{32}$$

with control vector u, an endogenous state vector x that is governed by the law of motion $g(x, u)$, and produces return function $r(x, u)$. The stochastic exogenous state vector is drawn from the set $\mathcal{S}$, with probability distribution $\lambda(s'|s)$ which is conditional on the current exogenous state vector s. Under one of several possible sets of regularity conditions[12] the solution to this dynamic programming problem is a unique policy function $u = h(x, s)$.

Regularity conditions can ensure a unique policy function for an individual facing exogenous processes: Robinson Crusoe faces no multiplicity when pricing his papaya trees (Lucas, 1978). But the same conditions may not give uniqueness in a decentralized economy, where individuals treat endogenous aggregates as exogenous random variables.

To see why, consider the decentralized dynamic programming problem considered by Prescott and Mehra (1980), in their original study of recursive competitive equilibria with many homogeneous agents. Again, let x denote an agent's endogenous state, but let z denote an exogenous stochastic state that affects all agents. In such a problem, agents take others' states as given, so from any agent's perspective, the exogenous state is the vector $s = \{\underline{x}, z\}$, where $\underline{x}$ is the vector of other agents' endogenous states. $f(\underline{x}, z) = \underline{x}'$ is the endogenous evolution rule for the average state $\underline{x}$. Prescott and Mehra write (albeit with different notation and less generality) an individual's dynamic programming problem as

$$V(x, \{\underline{x}, z\}) = \max_{u} r(x, z, u) + \beta \int_{z' \in \mathcal{Z}} V(g(x, z, u), \{f(\underline{x}, z), z'\}) d\lambda(z'|z) \tag{33}$$

which, under regularity conditions, admits a unique policy function $u = h(x, \{\underline{x}, z\})$. Agents are homogeneous, so when symmetry is imposed, this gives the symmetric policy function $\underline{u} = h(\underline{x}, \{\underline{x}, z\})$. A *Recursive Competitive Equilibrium* is characterized by a policy function and evolution rule consistent with one another, satisfying

$$f(\underline{x}, z) = g(\underline{x}, z, h(\underline{x}, \{\underline{x}, z\}))$$

To show that their recursive competitive equilibrium is unique, Prescott and Mehra point out that the Welfare Theorems hold given their assumption that neither $r(x, z, u)$ nor $g(x, z, u)$ depend on $\underline{x}$. Thus it suffices to solve the social planner's problem with symmetric Pareto weights:

$$W(\underline{x}, z) = \max_{\underline{u}} r(\underline{x}, z, \underline{u}) + \beta \int_{z' \in \mathcal{Z}} W(g(\underline{x}, z, \underline{u}), z') d\lambda(z'|z) \tag{34}$$

[12] See Stokey, Lucas, and Prescott (1989) for examples of such regularity conditions.

with the solution being the policy function $\underline{u} = m(\underline{x}, z)$. The solution to the social planner's problem is unique under the same regularity assumptions that ensure the individual's problem (33) is unique, and moreover the social planner's policy function is exactly the symmetric policy function:

$$m(\underline{x}, z) = h(\underline{x}, \{\underline{x}, z\})$$

By construction, the solution $m(\underline{x}, z)$ is consistent with the law of motion for the aggregate state $f(\underline{x}, z)$, so Prescott and Mehra conclude that there exists a unique recursive competitive equilibrium.

Except, writing the individual's dynamic programming problem as equation (33) does not immediately follow from the structure of the problem. Doing so implicitly assumes the nescient equilibrium! First, generalize the agent's exogenous state to be $s = \{\vec{\underline{x}}, z\}$, where $\vec{\underline{x}}$ is a (possibly infinite) vector of current and past states.[13] Correctly applying the individual's exogenous state s to the generic Bellman equation (32) yields:

$$V(x, \{\vec{\underline{x}}, z\}) = \max_u r(x, z, u) + \beta \int_{\{\vec{\underline{x}}', z'\} \in \mathcal{S}} V(g(x, z, u\}, u), \{\vec{\underline{x}}', z'\}) dF(\{\vec{\underline{x}}', z'\} | \{\vec{\underline{x}}, z\}) \tag{35}$$

where $F(\{\vec{\underline{x}}', z'\} | \{\vec{\underline{x}}, z\})$ is a conditional probability distribution.

Therein lies the problem. The probability distribution $F(\{\vec{\underline{x}}', z'\} | \{\vec{\underline{x}}, z\})$ is endogenous; it depends on the dynamics of $\vec{\underline{x}}$, which are determined in equilibrium. Prescott and Mehra's formulation presumes that agents have the nescient information set, because they forecast with the exogenous probability distribution $\lambda(z'|z)$. Their argument for uniqueness depends on the fact that the social planner's problem (equation 34) has exactly the same conditional probability distribution as their decentralized problem (equation 33). But it may not. The nescient solution that solves the social planner's problem is one valid equilibrium, but there can be more.

The stationary equilibria calculated in Section 3.3 serve as examples. The Bellman equation is just the asset pricing equation (3), with no endogenous controls or states. The exogenous state vector is the history of prices and shocks, $\{p_{t-j}, \varepsilon_{t-j}\}_{j=0}^{\infty} = \Omega_t$. The return function is the current dividend x_t. And the conditional probability distribution $F(\{\vec{\underline{x}}', z'\} | \{\vec{\underline{x}}, z\})$ is determined by the information basis w_t and the equilibrium price process given by Theorem 3.

[13] Considering the state $s = \{\underline{x}, z\}$ presumes that $\underline{x}$ is first order Markov, which need not be true. For example, the nescient equilibrium of the AR(1) asset pricing model in Section 3.3 was first order Markov, but equilibria for other bases were not.