Revisiting the Forecasts of Others*

Ryan Chahrour
Cornell University

Kyle Jurado
Duke University

November 8, 2024

Abstract

In macroeconomic models with dispersed information, agents have an incentive to learn from endogenous variables, which themselves depend on the forecasts of others. This paper revisits the model of Townsend (1983) to characterize how this mechanism affects equilibrium dynamics. The first part of the paper simplifies, revises, and extends past results about situations when prices are fully revealing. The second part explains that full revelation does not occur in the original model and proves that the equilibrium state vector is infinite-dimensional. It also provides a new numerical solution procedure for such cases, which operates entirely in the frequency domain.

JEL classification: E21, D83

Keywords: incomplete information, endogenous signals, higher-order expectations.

*Chahrour: Department of Economics, Cornell University, 404 Uris Hall, Ithaca, NY 14853 (e-mail: `ryan.chahrour@cornell.edu`); Jurado: Department of Economics, Duke University, 419 Chapel Drive, Durham, NC 27708 (e-mail: `kyle.jurado@duke.edu`). We would like to thank Bartosz Mackowiak for his helpful discussion of this paper.

1 Introduction

When agents have heterogeneous and imperfect information about the state of the economy, they each have an incentive to learn from their observations of endogenous aggregate variables when making forecasts. But because these variables are endogenous, their information content simultaneously depends on the solution to the forecasting problems simultaneously being solved by other agents in the economy. This mechanism has proven to be both interesting and challenging for economists to incorporate into their models. Interesting both because it can alter the way that fundamental shocks propagate through the economy and because it opens the door for non-fundamental shocks to expectations to have real consequences, but challenging because it can introduce technical difficulties for standard solution procedures. In the existing literature, the model of Townsend (1983) has played an important role as an early dynamic formalization of this mechanism, and as a laboratory in which to explore its implications.

The purpose of this paper is to revisit the Townsend model to simplify, revise, and extend existing theoretical results about it in the large (and growing) subsequent literature. The first part of the paper shows that an aggregate price index can reveal so much information about the state of the economy that uncertainty about other firms' forecasts plays no role in affecting the equilibrium dynamics. Existing proofs of this result appear in the literature, but with disadvantages, in that they are either less general, unnecessarily roundabout, or incorrect. Furthermore, this part proves that the revealing equilibrium is unique, which is more difficult to establish, and has so far proven elusive. It then describes how this collection of results extends to perturbed versions of the baseline model, including versions with persistent idiosyncratic shocks and structural heterogeneity across sectors.

The second part of the paper discusses the version of the model originally analyzed by Townsend, in which observations of the aggregate price index are not perfect, but are contaminated by independent noise. The main contribution in this part is an impossibility result, which says it is impossible to represent the equilibrium dynamics with a finite number of state variables. An equivalent way to say this is that, even though the endogenous processes are all stationary, they do not have autoregressive moving average (ARMA) representations. This formally confirms Townsend's original conjecture that the infinite regress of higher-order beliefs in this model leads to an infinite state problem, despite evidence to the contrary from the existing literature.

The fact that the state is infinite-dimensional poses a challenge for using standard Kalman filtering formulas to compute the equilibrium, and this paper presents a new numerical procedure to compute the equilibrium in models of this type by iterating on the equilibrium fixed point equation in the frequency domain. This procedure is used to compare the predictions of the model with and without learning from endogenous variables in a numerical example. This example shows that a natural modification of the Townsend model in which firms receive a noisy signal of the exogenous aggregate demand shock instead of the endogenous aggregate price index makes very similar predictions, while avoiding the complications that arise from having an endogenous signal. Of course, this finding is model-specific, and the additional discipline and different counterfactual predictions of the endogenous signal model are still reasons why one might prefer this formulation.

The approach of this paper is to focus attention narrowly on the Townsend model rather than to state results over a more abstract class of models. The cost of this approach is that the specific results in this paper cannot be directly applied to other models without modification. However, the benefit is that by restricting attention to a particular model, it is possible to take results farther and make them more concrete. By working through each step of the analysis in as much detail as possible, it will hopefully be easier to understand both the results themselves and how to use the same steps to prove similar results in other models.

The paper is most closely connected to a series of papers that directly analyze the Townsend model. Marcet and Sargent (1989), Sec. III, use a least-squares learning algorithm to compute the equilibrium of the model numerically, under the restriction that agents' perceived laws of motion are first-order vector autoregressions. Sargent (1991) extends this algorithm by allowing agents to fit vector ARMA models, and claims that by doing so it is possible to formulate the equilibrium as the fixed point of a finite-dimensional operator. Taub (1989), Sec. 5, explains that full revelation can obtain in a model similar to Townsend's with a large number of agents and perfect observation of aggregate capital. Kasa (2000) seeks to derive the closed-form solution to the Townsend model without assuming that the state of the economy is fully revealed after a finite number of periods. Pearlman and Sargent (2005) apply the methodology proposed by Pearlman et al. (1986) to show that prices can fully reveal demand shocks in a two-sector version of the model. Points of connection with these papers are discussed as they arise in the analysis below.

Beyond the Townsend model, this paper is also related to the broader literature on learning from endogenous signals. Many early models of this mechanism assume that learning only lasts for one period, as in the static models of Grossman (1976), Kreps (1977), Grossman and Stiglitz (1980), and Hellwig (1980), or the dynamic models of Lucas (1972), King (1982), and Kimbrough (1984). Papers that follow Townsend in allowing learning from endogenous variables to last multiple periods include Chari (1979), Futia (1981), Singleton (1987), He and Wang (1995), Bacchetta and Van Wincoop (2006), Hellwig and Venkateswaran (2009), Bernhardt et al. (2010), Makarov and Rytchkov (2012), Kasa et al. (2014), Melosi (2016), Nimark (2017), Rondina and Walker (2021), Acharya et al. (2021), Sec. 5 of Miao et al. (2021), Han et al. (2022), Adams (2022), Rondina and Walker (2023), Sec. 5 of Huo and Takayama (2023), and Huo and Pedroni (2023). Another part of the literature also emphasizes the importance of higher-order beliefs, but in models with no learning from endogenous variables. Examples include Morris and Shin (2002), Woodford (2003), Lorenzoni (2009), Angeletos and La'O (2013), and Nimark (2014). Angeletos and Lian (2016) provides a detailed review of the literature on dispersed information.

To outline the paper, Section 2 describes the Townsend model and defines the rational expectations equilibrium up to a specification of agents' information sets. Section 3 characterizes situations in which a single index of prices reveals enough information for firms to act as if all private information was commonly known. Section 4 analyzes the case when prices are observed only with error and proves that the state vector becomes infinite-dimensional. Section 5 concludes.

2 Townsend model

This section describes the model of Townsend (1983). It is a multi-sector version of the Lucas and Prescott (1971) model of firm investment under uncertainty, where the only interconnection between sectors arises through the structure of demand. The description provided here differs from the original in explicitly deriving the system of linear equilibrium conditions as approximations from a nonlinear model.

The economy is made up of n sectors, each of which has a representative firm. At each point in time, the firm in sector i chooses a contingent plan for investment from that time forward, so as to maximize expected discounted cash flows. From the perspective of time $t = 0$, the firm chooses I_{it} for all $t \geq 0$ so as to maximize

$$E_{i0} \sum_{t=0}^{\infty} \beta^t \left[P_{it} Y_{it} - I_{it} \left(1 + \Phi \left(\frac{I_{it}}{K_{it}} \right) \right) \right],$$

where E_{i0} denotes the expectations of the firm in sector i as of time $t = 0$, P_{it} is the sectoral price of output, Y_{it} is the output of the firm, I_{it} is gross investment expenditure on new capital goods, $\beta \in (0, 1)$ is an intertemporal discount factor, and Φ is a strictly increasing and convex adjustment cost function satisfying $\Phi(0) = 0$ and $2\Phi'(\cdot) + \alpha\Phi''(\cdot) > 0$, as in Abel and Blanchard (1983). The representative firm assumption implies that the notation Y_{it}, I_{it}, and K_{it} can be used interchangeably for sector-level and firm-level variables; the same is true for the operator E_{it}.

The firm's maximization problem is subject to the production technology $Y_{it} = F(K_{it})$, where F is strictly increasing and concave, the capital accumulation equation $K_{i,t+1} - K_{it} = I_{it} - \delta K_{it}$, where $\delta \in (0, 1)$ is the depreciation rate of the capital stock, and the long-run constraint $\lim_{t\to\infty} \beta^t E_{i0} K_{it} \geq 0$. The timing convention adopted here is that output is produced using the stock of capital that was determined one period in advance.

Up to a log-linear approximation, the optimal evolution of the capital stock in sector i can be described by the equation

$$f_2(k_{i,t+1} - k_{it}) = \beta E_{it}[f_0 p_{i,t+1} - f_1 k_{i,t+1} + f_2(k_{i,t+2} - k_{i,t+1})], \tag{1}$$

where $k_{it} \equiv \ln(K_{it}/K_i)$ and $p_{it} \equiv \ln(P_{it}/P_i)$ denote the percent deviation of capital and price from their steady state values $K_i > 0$ and $P_i > 0$, $f_0 \equiv P_i F'(K_i) > 0$, $f_1 \equiv -P_i F''(K_i) K_i \geq 0$, and $f_2 \equiv 2\Phi'(\delta) + \delta\Phi''(\delta) > 0$. The steady state values are the values to which the variables in the model converge in the absence of any exogenous disturbances, and all subjective expectations are correct. The analysis abstracts from trend growth, which is why the steady-state values of capital and the price of output are constant.

The price of output in each sector is determined in equilibrium, which requires a specification of demand. This is done by introducing a demand schedule for the output of each sector of the form $P_{it} = D(Y_{it}, U_{it})$, where D is strictly decreasing in Y_{it} and strictly increasing in U_{it}, which is an exogenous random variable. Importantly, U_{it} is not independent across sectors. Exogenous shifts to demand in sector i are at least partly correlated with shifts to demand in other sectors. This correlation creates a physical link between sectors, and provides an incentive for firms in one sector to extract information from variables in other sectors about their own demand.[1]

Up to a log-linear approximation around the steady state, the demand schedule in sector i can be described by the equation

$$p_{it} = -b_1 y_{it} + u_{it}, \tag{2}$$

where $y_{it} \equiv \ln(Y_{it}/Y_i)$ denotes the percent deviation of output from its steady-state value $Y_i > 0$, $u_{it} \equiv D_U(Y_i, 0)/D(Y_i, 0)U_{it}$ is proportional to the deviation of U_{it} from its steady-state value $U_i = 0$, and $b_1 \equiv -D_Y(Y_i, 0)K_i/D(Y_i, 0)^2 > 0$. In addition, it is assumed that $D(Y_i, 0) = K_i/Y_i > 0$, so the production function can be written in log-linear approximate form as

$$y_{it} = f_0 k_{it}. \tag{3}$$

The exogenous component of demand, u_{it}, is represented as the sum of a persistent economy-wide component θ_t and a transitory idiosyncratic component ε_{it},

$$u_{it} = \theta_t + \sigma_\varepsilon \varepsilon_{it}, \qquad \theta_t = \rho\theta_{t-1} + \sigma_v v_t, \tag{4}$$

where $\rho \in (0, 1)$, $\sigma_\varepsilon, \sigma_v > 0$, and the random variables $v_t, \varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt}$ are jointly Gaussian, mutually uncorrelated and uncorrelated over time, with mean zero and unit variance.[2] Note that, by the law of large numbers, $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \varepsilon_{it} = 0$.

The system of equations (1), (2), (3), and (4) describes the equilibrium in the economy at each point in time, up to a specification of expectations.[3] It represents a "temporary equilibrium" of the type discussed by Woodford (2013), and is compatible with a range of different assumptions regarding how expectations are formed, provided that these expectations satisfy standard probability laws (e.g. $E_{it} = E_{it}E_{i,t+1}$). The focus in this paper is on rational expectations equilibria, which means that expectations are formed rationally as a function of the variables firms observe when they make their investment decisions.

Letting s_{it} denote the observation vector of the representative firm in sector i

[1]Another reason firms might find information about other islands informative is the presence of strategic complementarity, as in Woodford (2003).

[2]Gaussianity can be dispensed with if the expectations in (1) are interpreted as linear projections.

[3]Note that the reduced-form parameters in (1), (2), and (3) exactly match the original notation of Townsend (1983). This shows that the model considered in that paper can be interpreted as a linear-quadratic approximation of the nonlinear model described here.

at time t, its information set at that time is given by the information generated by the current and past history of this observation vector, $s_i^t \equiv (s_{it}, s_{i,t-1}, \dots)$, so that $E_{it} = E(\cdot|s_i^t)$. As usual, this assumes that information is retained over time. Any variables that are either directly chosen by the firm at time t or are functions of them, such as $k_{i,t+1}$ and $y_{i,t+1}$, must be measurable with respect to s_i^t, and so are always contained in the firm's time-t information set. Other endogenous variables from sector i that are not directly chosen by the firm may or may not be contained in its information set, depending on the specification of s_{it}. For example, (2) implies that p_{it} will be contained in the firm's time-t information set if s_{it} includes u_{it}.

The distinctive feature of the Townsend model is that firms' information sets can depend on endogenous variables from other sectors. To the extent that equilibrium prices are correlated across sectors, due to correlated demand, the firm wishes to extract whatever information from these variables is helpful for predicting prices in its own sector. But because the variables are endogenous with respect to the economy as a whole, their information content depends on the solution to the signal extraction problems simultaneously being solved by firms in other sectors. This feature would not be present if the observation vector s_{it} consisted only of exogenous variables.

It is now possible to define a rational expectations equilibrium in this economy, up to a specification of the observation vectors $s_{1t}, s_{2t}, \dots, s_{nt}$. Let us collect all exogenous random variables, including $v_t, \varepsilon_{1,t}, \varepsilon_{2t}, \dots, \varepsilon_{nt}$, and any exogenous random variables introduced in the specification of s_{it}, into the single vector ξ_t.[4] Attention will be restricted to stationary equilibria, in which the process $\{\xi_t\}$ has a stationary structure and extends back into the infinite past, and all endogenous variables are time-invariant linear functions of the history $\xi^t \equiv (\xi_t, \xi_{t-1}, \dots)$. This abstracts from transitional dynamics, and amounts to analyzing only the limiting stationary distribution of the economy. The linearity restriction reflects the fact that we are interested in characterizing the equilibrium dynamics only up to a (log-)linear approximation.

Definition 1. A *rational expectations equilibrium* (REE) is a collection of covariance stationary processes $\{y_{it}, k_{it}, p_{it}\}$ for each sector that satisfy (1), (2), and (3), given an exogenous demand process $\{u_{it}\}$ that satisfies (4) and an observation vector s_{it} that is a time-invariant linear function of ξ^t.

[4]These may include non-fundamental noise or sunspot variables, or variables containing news about future fundamental disturbances.

3 Information revelation

This section proves that uncertainty about higher-order beliefs plays no role in equilibrium dynamics when, in addition to economic conditions in its own sector, each firm is able to observe the economy-wide average output price. The reason is that, in equilibrium, the average price reveals the average demand shock, which is a sufficient statistic of the demand shocks in all sectors. By observing the average price, each firm is able to implement the same state-contingent plan that it would choose if it were able to observe all the demand shocks directly. The existing literature contains partial versions of this result, which establish only that an information revealing equilibrium of this type exists in certain special cases. The purpose of this section is to simplify and extend those results, and then to present new results regarding the more difficult question of whether this equilibrium is unique.

Before analyzing the equilibrium in which firms must learn from endogenous variables, it is necessary to introduce a type of equilibrium originally proposed by Radner (1979), in which private information about demand is shared by firms in all sectors.[5]

Definition 2. A *full communication equilibrium* (FCE) is a REE with $s_{it} = u_t \equiv (u_{1t}, u_{2t}, \ldots, u_{nt})$ for all i and t.

In a FCE, all firms in the economy have the same information at each point in time, which consists of the history $u^t \equiv (u_t, u_{t-1}, \ldots)$. This implies that output and capital in all sectors are common knowledge. It also implies that firms' forecasts of all variables are the same. Higher-order uncertainty plays no role in this equilibrium because each firm knows the forecasts of all other firms. In a FCE, however, firms still have imperfect information about the underlying disturbances v_t and $\varepsilon_t \equiv (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt})$, (at least for finite n), and so about the realization of ξ_t. Therefore, it is possible to distinguish a FCE from a "full information equilibrium," in which the history of all exogenous disturbances is common knowledge; i.e. $s_{it} = \xi_t$.[6]

The first result is the closed-form solution to the FCE. The optimal capital choice of firm i is a second-order autoregressive transformation of the average demand shock.

[5]Sometimes the FCE is described as a different equilibrium concept from the REE; but it is equally possible to view it as a REE with a particular specification of information, as is done here. Sometimes this equilibrium is also referred to as a "pooling equilibrium."

[6]According to this terminology, the full communication equilibrium approaches the full information equilibrium as $n \to \infty$, since then all firms can perfectly infer ξ_t at each point in time.

Proposition 1. *The FCE exists and is unique. Moreover, in this equilibrium,*

$$k_{i,t+1} = \frac{\omega}{(1-\lambda L)(1-\phi L)}\bar{u}_t \tag{5}$$

for all i and t, where L denotes the lag operator, $\bar{u}_t \equiv \frac{1}{n}\sum_{i=1}^n u_{it}$, $\lambda \in (0,1)$ solves $\lambda^2 - (1+\beta^{-1} + (b_1 f_0^2 + f_1)/f_2)\lambda + \beta^{-1} = 0$, $\phi \in (0,\rho)$ solves $\rho\sigma_\varepsilon^2\phi^2 - (\sigma_\varepsilon^2(1+\rho^2) + n\sigma_v^2)\phi + \rho\sigma_\varepsilon^2 = 0$, and $\omega \equiv \frac{f_0\beta\lambda(\rho-\phi)}{f_2(1-\beta\lambda\rho)} > 0$.

The details of the proof are in Appendix A, but it is helpful to provide a brief sketch here. Using the demand schedule (2) to substitute out the price from the capital optimality condition (1), the equilibrium path of capital in sector i must evolve according to the equation

$$k_{i,t+1} = \lambda k_{it} + \frac{f_0}{f_2}\sum_{j=1}^{\infty}(\beta\lambda)^j E_{it}u_{i,t+j}, \tag{6}$$

where λ has the definition stated in the proposition. This says that capital in sector i depends only on forecasts of future demand shocks in sector i. The orthogonal projection theorem implies that the conditional expectation $E_{it}u_{i,t+j} \equiv E(u_{i,t+j}|u^t)$ exists and is unique for all i, t, and j, so the equilibrium capital path exists and is unique as well. The closed-form solution in (5) is obtained by explicitly computing these expectations using the structure of the demand process in (4).

The second result applies to an economy with dispersed information, in which firms are not able to directly observe the history of demand shocks in all sectors. Each firm still observes the demand shock in its own sector, but now in addition can only observe the economy-wide average output price. What can be shown in this case is that the FCE paths from Proposition 1 are a REE in this economy.

Proposition 2. *Consider an economy with $s_{it} = (u_{it}, \bar{p}_t)$ for all i and t, where $\bar{p}_t \equiv \frac{1}{n}\sum_{i=1}^n p_{it}$. The FCE paths of $\{y_{it}, k_{it}, p_{it}\}$ are a REE in this economy.*

The proof of this result is straightforward. Equation (5) indicates that, in the FCE, capital in each sector depends only on the history of average demand shocks, $\bar{u}_t$. By the demand curve (2), the average price also depends only on the history of average demand shocks; i.e.

$$\bar{p}_t = \left[1 - \frac{b_1 f_0 \omega L}{(1-\lambda L)(1-\phi L)}\right]\bar{u}_t. \tag{7}$$

Using the closed-form expressions for ω, ϕ, and λ, we show in Appendix A that the operator on the right side of (7) is invertible for any admissible values of the model parameters. This implies that the history of average prices contains the same information as the history of average demand shocks, so each firm will implement the same state-contingent plan that is optimal under full communication. In principle, the operator in (7) could fail to be invertible, if the effect of past demand shocks on prices is so large relative to the effect of current demand shocks on prices that firms are unable to identify the current average demand shock even with the entire history of average prices. What Proposition 2 shows is that this is not possible in this version of the Townsend model.

While Proposition 2 says that the FCE paths of $\{y_{it,}, k_{it}, p_{it}\}$ are a REE in the dispersed information economy, this does not imply that the two equilibria are identical. Firms still have less information in the dispersed information economy. For example, they only have imperfect information about output and prices in sectors other than their own. In principle, observations of cross-sector forecasts of these variables would be able to distinguish between these two equilibria. Instead, what Proposition 2 says is that there exist equilibrium paths of $\{y_{it,}, k_{it}, p_{it}\}$ which are the same "as if" firms had this additional information. The average output price aggregates all the relevant information necessary to optimally predict their own future prices.

For the special case with $n = 2$, a different proof of this result is provided in Sec. 5.2 of Pearlman and Sargent (2005). The strategy in that paper is to apply the method developed by Pearlman et al. (1986) to show by brute force computation that perceived laws of motion coincide with actual laws of motion when the perceived laws of motion for each firm are the ones from the FCE. What is shown here is that it is possible to avoid that computation by instead checking that the operator in (7) is invertible. This is both simpler and more intuitive, because it shows that the reason observing output prices in both sectors is sufficient for implementing FCE plans is because they can be used to compute the average price, the history of which can be used to infer the history of average demand shocks.

Sec. 3 of Kasa (2000) also provides a proof of this result with $n = 2$, based on computing the closed-form solution of the model and inspecting its properties. The problem is that the closed-form solution provided there, described in his Proposition 3.1.2, is not correct. The reason for this is that the frequency-domain procedure used to compute the solution assumes that each of the three variables in the observation vector contains some independent information. The observation vector in that paper is (u_{it}, p_{1t}, p_{2t}), which in the case of $n = 2$ is informationally equivalent to $\tilde{s}_{it} = (u_{it}, p_{it}, \bar{p}_t)$. This vector is three dimensional, but only contains two independent sources of information in the FCE.[7] To see this, observe that (2) and (5) imply

$$p_{it} = \frac{-b_1 f_0 \omega L}{(1-\lambda L)(1-\phi L)} \bar{u}_t + u_{it}.$$

Proposition 2 proves that $\bar{p}^t$ and $\bar{u}^t$ contain the same information, so this equation shows that p_{it} is a function of $(u_i^t, \bar{p}^t)$, and so contains no additional information.[8]

While Proposition 2 improves upon existing proofs of this result in the literature, it says nothing about whether the equilibrium described there is unique. Could there perhaps exist other rational expectations equilibria in which the paths of $\{y_{it,}, k_{it}, p_{it}\}$ differ from their FCE paths? This question is substantially more difficult to answer, and so far there are no results about it in the existing literature.

The following proposition says that the equilibrium from Proposition 2 is the unique symmetric REE. The notion of symmetry involved is that all firms have the same policy functions mapping information sets into actions. For example, $k_{i,t+1} = B(L)s_{it}$, where $B(L)$ is the same for all i. This does not require all firms to have the same information, of course, because the realizations of s_{it} can differ across sectors.

Proposition 3. *In any symmetric REE of the economy from Proposition 2, the paths of $\{y_{it}, k_{it}, p_{it}\}$ are the same as in the FCE.*

While the details of the proof are somewhat involved, the basic structure is straightforward, and consists of four main steps. The first is to prove that in any symmetric REE, there is a mapping of the form

$$\bar{p}_t = A(L)\bar{u}_t \tag{8}$$

from average demand shocks to average prices, where $A(L)$ is a scalar operator which is one-sided into the past. Equation (7) shows that a mapping of this form exists in the FCE; the point here is to show such a mapping holds in any symmetric REE.

The remaining three steps involve showing that in any REE in which a relation of the form (8) holds, the operator $A(L)$ must be invertible, so the history of average

[7]More formally: the three-dimensional process $\{\tilde{s}_{it}\}$ has rank two; cf. Sec. 1.9 of Rozanov (1967).
[8]Alternatively, it shows that u_{it} is redundant given $(p_i^t, \bar{p}^t)$.

prices and average demand shocks contain the same information. The second step uses (8) and the law of motion of u_{it} in (4) to find the Wold representation of the observation process, $s_{it} = \Gamma(L)w_{it}$, where w_{it} is the one-step-ahead innovation in s_{it}, and the operator $\Gamma(L)$ depends on $A(L)$. This operator is needed to compute firms' optimal forecasts of future demand. Usually, these forecasts are computed using the Kalman filter. But because $A(L)$ is arbitrary, the older Wiener-Kolmogorov filter must be used instead.[9] The third step uses $\Gamma(L)$ and the structural equations of the model to find the fixed point equation $A(L) = T[A(L)]$ in closed form. The fourth step shows that any operator $A(L)$ satisfying this fixed point equation must be invertible, which completes the proof.

A first remark regarding Proposition 3 is that it proves that this model has no equilibria in which endogenous variables respond to sunspots (random variables independent of the fundamental disturbances v_t and ε_t). For example, if we conjecture, following Acharya et al. (2021), that $\bar{p}_t = A(L)\bar{u}_t + C(L)\eta_t$, where η_t is a vector of sunspot shocks, is it possible that $C(L)$ is non-zero? The answer is no, because (8) is a *necessary condition* of equilibrium, and implies $C(L) = 0$. To see why (8) is necessary, notice that $k_{i,t+1}$ is measurable with respect to $s_i^t = (u_i^t, \bar{p}^t)$, so

$$k_{i,t+1} = B_u(L)u_{it} + B_p(L)\bar{p}_t$$

for some one-sided operators $B_u(L)$ and $B_p(L)$, which don't depend on i by symmetry. Substituting this into the demand curve (2) and averaging across i yields

$$\bar{p}_t = (1 - bf_0 B_u(L)L)\bar{u}_t - b_1 f_0 B_p(L)L\bar{p}_t.$$

Solving for $\bar{p}_t$ implies (8), with

$$A(L) = \frac{1 - b_1 f_0 B_u(L)L}{1 + b_1 f_0 B_p(L)L},$$

which must be one-sided because, by definition, $\bar{p}_t$ cannot be correlated with future disturbances in any REE.

Continuing this first remark, it might be thought that the necessity of (8) is sensitive to the assumption that $s_{it} = (u_{it}, \bar{p}_t)$, which only allows firms to condition their actions on sunspots *indirectly*, through the endogenous average price. What would

[9] Whittle (1983) is a standard reference on these two approaches to filtering.

happen in an alternative economy in which $s_{it} = (u_{it}, \bar{p}_t, \eta_t)$? Would the fact that firms in this economy can *directly* condition their actions on sunspots make it possible to sustain an equilibrium in which these sunspots affect equilibrium outcomes? The answer is again no. Firms understand that their own future demand shocks are independent of sunspots; when the sunspots are directly observed, rationality requires firms' demand shock forecasts to be independent of them as well. More formally, if $\hat{p}_t$ denotes the residual from a projection of $\bar{p}_t$ on $\eta^t = (\eta_t, \eta_{t-1}, \dots)$, then $E_{it}u_{i,t+j} = E(u_{i,t+j}|u_i^t, \bar{p}^t, \eta^t) = E(u_{i,t+j}|u_i^t, \hat{p}^t, \eta^t) = E(u_{i,t+j}|u_i^t, \hat{p}^t)$, which is independent of η^t. The demand curve (2) and the policy function (6) imply

$$\bar{p}_t = \bar{u}_t - \frac{b_1 f_0^2}{f_2} \frac{L}{1 - \lambda L} \sum_{j=1}^{\infty} (\beta\lambda)^j \bar{E}_t u_{i,t+j}, \tag{9}$$

so $\bar{p}_t$ must also be independent of η^t. This implies that sunspots cannot affect any of the endogenous variables in this economy either. Therefore, it is without loss of generality in Proposition 3 to exclude sunspot variables from the observation vector.

A second remark regarding Proposition 3 is that it applies only to *symmetric* equilibria. While this is required for the proof, there is nevertheless no positive reason to think that other non-symmetric equilibria exist which differ from the FCE. In fact, as it has perhaps been possible to infer from the discussion so far, the information-revealing properties of the average price imply an even starker uniqueness result, which does not require symmetry.

Proposition 4. *Consider an economy with $s_{it} = \bar{p}_t$ for all i and t. The FCE paths of $\{y_{it}, k_{it}, p_{it}\}$ constitute the unique REE in this economy.*

This result says that the average price in fact reveals so much information that once firms observe this, they do not need to observe any other information about demand in their own sector to implement the same state-contingent plan that would be optimal with full communication. The reason is that, as shown in equation (5), the optimal evolution of capital in the full communication equilibrium only depends on the history of average demand shocks. Since these are fully revealed by the average price, this information is sufficient for all firms to implement their optimal actions, regardless of whether they also observe other prices or quantities in their own sector.

Why does Proposition 4 imply that firms do not need to rely on their own sector-specific information to implement the optimal plan under full communication? The result is due to the assumption that the idiosyncratic component of demand is purely transitory, together with the one period time-to-build assumption in production. The optimal choice of capital today, to be used in production tomorrow, depends on forecasts of demand shocks from tomorrow out into the future, as shown in (6). Since today's idiosyncratic shock is purely transitory, it is only necessary to forecast the common component of demand; $E_{it}u_{i,t+j} = E_{it}\theta_{t+j}$. And since u_{it} does not contain any more information about this common component beyond what is contained in $\bar{u}_t$, it is therefore unnecessary to respond to it.

From this discussion, it is not difficult to see that Proposition 4 relies heavily on the assumption that common disturbances to demand are persistent, while idiosyncratic disturbances are not. While this assumption is a common one, and has been followed in much of the subsequent literature, there is no economic rationale for imposing it. Sector-specific variation in demand could be as or even more persistent than common variation in demand. The next sub-section considers this possibility in more detail.

3.1 Persistent idiosyncratic shocks

When the idiosyncratic component of demand is persistent, sector-specific information is useful for predicting future demand. In this case, the stark result of Proposition 4, that observing the average price alone is sufficient to implement the full communication plan, no longer holds. However, this subsection shows that generalizations of Propositions 1, 2, and 3 hold under this perturbation of the baseline model.

To introduce persistence in the idiosyncratic component of demand, generalize the law of motion (4) to

$$u_{it} = \theta_t + z_{it}, \qquad \theta_t = \rho_\theta \theta_{t-1} + \sigma_v v_t, \qquad z_{it} = \rho_z z_{i,t-1} + \sigma_\varepsilon \varepsilon_{it}, \tag{10}$$

where $\rho_\theta, \rho_z \in (0,1)$ and $\sigma_\varepsilon, \sigma_v > 0$. The new parameter ρ_z controls the persistence of the idiosyncratic component. Continue to assume that the random variables $v_t, \varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt}$ are jointly Gaussian, mutually uncorrelated and uncorrelated over time, with mean zero and unit variance.

The first result generalizes Proposition 1. It shows that when idiosyncratic shocks are persistent, the optimal choice of capital now depends both on $\bar{u}_t$ and u_{it}.

Proposition 5. *Consider an economy in which the demand process $\{u_{it}\}$ satisfies*

(10). *The FCE exists and is unique. Moreover, in this equilibrium,*

$$k_{i,t+1} = \frac{\omega_\theta(1-\rho_z L)}{(1-\lambda L)(1-\phi L)}\bar{u}_t + \frac{\omega_z}{1-\lambda L}u_{it} \tag{11}$$

for all i and t, where L denotes the lag operator, $\bar{u}_t \equiv \frac{1}{n}\sum_{i=1}^n u_{it}$, $\lambda \in (0,1)$ is defined as in Proposition 1, $\phi \in (0,1)$ solves $(\rho_\theta\sigma_\varepsilon^2 + \rho_z n\sigma_v^2)\phi^2 - (\sigma_\varepsilon^2(1+\rho_\theta^2) + n\sigma_v^2(1+\rho_z^2))\,\phi + (\rho_\theta\sigma_\varepsilon^2 + \rho_z n\sigma_v^2) = 0$, $\omega_\theta \equiv \frac{f_0\beta\lambda(\rho_\theta-\phi)}{f_2(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)}$, and $\omega_z \equiv \frac{f_0\beta\lambda\rho_z}{f_2(1-\beta\lambda\rho_z)} > 0$.

Equation (11) shows that the optimal capital choice places some weight on the history of both average and sector-specific demand shocks. The expression for ω_z shows that the weight on the latter is positive whenever $\rho_z > 0$, while the expression for ω_θ shows that in the special case when the common and idiosyncratic components of demand have exactly the same persistence, $\rho_\theta = \rho_z$, which implies that $\rho_\theta = \phi$, the optimal capital choice places no weight on the average demand shock. In this case, the full communication equilibrium coincides with the full information equilibrium, regardless of whether or not the number of sectors is finite. This is because common persistence implies that (10) reduces to $u_{it} = \rho u_{i,t-1} + \sigma_w w_{it}$, where $\rho \equiv \rho_\theta = \rho_z$, $\sigma_w^2 \equiv \sigma_v^2 + \sigma_\varepsilon^2$, and w_{it} is i.i.d. over time with mean zero and unit variance. Therefore, the current demand shock becomes a sufficient statistic for predicting future demand shocks: $E(u_{i,t+j}|\theta^t, \varepsilon^t) = E(u_{i,t+j}|u^t) = \rho^j u_{it}$ for all $j \geq 0$.

The second result generalizes Proposition 2. It shows that the FCE paths of the endogenous variables are always a REE in the dispersed information economy, even when idiosyncratic shocks are persistent.

Proposition 6. *Consider an economy in which the demand process $\{u_{it}\}$ satisfies (10) and $s_{it} = (u_{it}, \bar{p}_t)$ for all i and t, where $\bar{p}_t \equiv \frac{1}{n}\sum_{i=1}^n p_{it}$. The FCE paths of $\{y_{it}, k_{it}, p_{it}\}$ are a REE in this economy.*

The logic of the proof is the same as in Proposition 2. By subsituting the closed-form solution (11) for the FCE evolution of capital into the demand schedule (2) and averaging across sectors, it is possible to show that the average price depends on the history of average demand shocks in the following way,

$$\bar{p}_t = \left[1 - \frac{b_1 f_0 L}{(1-\lambda L)(1-\phi L)}\Big(\omega_\theta(1-\rho_z L) + \omega_z(1-\phi L)\Big)\right]\bar{u}_t \tag{12}$$

And, as in the proof of Proposition 2 it is possible to show that the operator on the right side of this equation is always invertible.

Propositions 5 and 6 imply that in the special case of common persistence, firms' private signals reveal enough information for them to implement their optimal plans under full information. This is consistent with Hellwig and Venkateswaran (2009), who show, in the context of new-Keysenian price setting models, that private endogenous information can reveal the firm's optimal price when common and idiosyncratic fundamentals have the same persistence. The result here is stronger, showing that the average price aggregates all the relevant private information from other sectors regardless of the relative persistence of common and idiosyncratic fundamentals.

The last proposition in this subsection extends Proposition 3 to the case of persistent idiosyncratic shocks, showing that the equilibrium from Proposition 6 is unique.

Proposition 7. *In any symmetric REE of the economy from Proposition 6, the paths of* $\{y_{it}, k_{it}, p_{it}\}$ *are the same as in the FCE.*

An implication of Propositions 5 and 7 is that in any equilibrium, endogenous public information is ignored when common and aggregate fundamentals have the same persistence, consistent with Proposition 1 of Taub (1989).

3.2 Structural heterogeneity

The results in this section so far have shown that in the Townsend model, the average economy-wide price has strong information revelation properties. A natural question is how much these results depend on the assumption that the sectors are completely symmetric, both with respect to supply and demand. To address this question, this section perturbs the baseline model by introducing different types of heterogeneity and exploring the extent to which previous results need to be modified.

The first modification is to relax the assumption that the structural parameters in equations (1), (2), and (3) are the same across sectors. The equations take the same form as before, but with all parameters explicitly indexed by i:

$$f_{2i}(k_{i,t+1} - k_{it}) = \beta_i E_{it}[f_{0i}p_{i,t+1} - f_{1i}k_{i,t+1} + f_{2i}(k_{i,t+2} - k_{i,t+1})] \tag{13}$$

$$p_{it} = -b_{1i}y_{it} + u_{it} \tag{14}$$

$$y_{it} = f_{0i}k_{it}. \tag{15}$$

The parameters continue to satisfy the inequalities $f_{0i} > 0$, $f_{1i} \geq 0$, $f_{2i} > 0$, and $b_{1i} > 0$ for all $i = 1, 2, \ldots, n$. Everything else about the economy and the definition of equilibrium remains the same as in Proposition 2. In this case, it is possible to prove the following result.

Proposition 8. *Consider an economy in which the REE paths of* $\{y_{it}, k_{it}, p_{it}\}$ *in each sector satisfy* (13), (14), *and* (15), *and the demand process* $\{u_{it}\}$ *satisfies* (4). *Propositions 1, 2, and 4 are true as stated, provided that* λ *and* ω *are replaced by* λ_i *and* ω_i, *where* $\lambda_i \in (0,1)$ *solves* $\lambda_i^2 - (1 + \beta_i^{-1} + (b_{1i} f_{0i}^2 + f_{1i})/f_{2i})\lambda_i + \beta_i^{-1} = 0$, *and* $\omega_i \equiv \frac{f_{0i}\beta_i\lambda_i(\rho-\phi)}{f_{2i}(1-\beta_i\lambda_i\rho)} > 0$.

This result says that, as long as the structure of demand shocks remains symmetric across sectors, other forms of heterogeneity do not not affect the information revelation properties of the average price. The intuition is that the symmetry of demand shocks implies that firms only need to obtain information about the average demand shock in order to implement their optimal state-contingent plans. In the FCE, the average price continues to be an invertible function of current and past average demand shocks, regardless of whether there is asymmetry across sectors in terms of their structural parameters. The only change relative to the results from the previous section is that proving that this mapping is invertible is somewhat more involved, as shown in Appendix A.

In the existing literature, Sec. 3 of Kasa (2000) suggests that structural asymmetries across sectors, such as in adjustment cost parameter, will prevent the full communication dynamics from being an equilibrium with partial information. The intuition provided there is that this type of asymmetry "jams the price signal," making it impossible to "posit symmetric responses in the two industries" to each of the two idiosyncratic shocks. The intuition that responses are no longer symmetric is correct, because the dependence of capital (and prices) on each of the idiosyncratic shocks depends on the parameters ω_i and λ_i, which can differ across sectors; e.g.

$$k_{i,t+1} = \frac{\omega_i}{(1-\lambda_i L)(1-\phi L)} \left(\theta_t + \frac{1}{n} \sum_{i=1}^{n} \sigma_\varepsilon \varepsilon_{it} \right).$$

However, what Proposition 1 demonstrates is that symmetric responses to idiosyncratic shocks is not a necessary condition for average prices to be a sufficient statistic for implementing full communication plans. This same conclusion is briefly mentioned in Sec. 5.2.1 of Pearlman and Sargent (2005), but only in terms of one type of asymmetry (adjustment costs) and only in a two-sector economy.

Now, consider the consequences of relaxing the assumption that demand shocks are symmetric across sectors. Specifically, suppose that (4) is generalized to allow differences both in the sensitivity of different sectors to the common component, and in the volatility of the idiosyncratic component,

$$u_{it} = \alpha_i \theta_t + \sigma_{\varepsilon i} \varepsilon_{it}, \qquad \theta_t = \rho \theta_{t-1} + \sigma_v v_t, \tag{16}$$

where $0 < \rho < 1$, $\sigma_{\varepsilon 1}, \sigma_{\varepsilon 2}, \ldots, \sigma_{\varepsilon n}, \sigma_v > 0$, and the random variables $v_t, \varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{nt}$ are jointly Gaussian, mutually uncorrelated and uncorrelated over time, with mean zero and unit variance.

Proposition 9. *Consider an economy in which the rational expectations paths of* $\{y_{it}, k_{it}, p_{it}\}$ *in each sector satisfy* (13), (14), *and* (15), *and the demand process* $\{u_{it}\}$ *satisfies* (16). *Propositions 1, 2, and 4 are true as stated, provided that* λ *is replaced with* λ_i *defined in Proposition 8,* ω *is replaced with* $\omega_i \equiv \alpha_i \frac{f_{0i}\beta_i\lambda_i(\rho-\phi)}{f_{2i}(1-\beta_i\lambda_i\rho)}$, σ_ε^2 *is defined by* $\frac{1}{\sigma_\varepsilon^2} \equiv \frac{1}{n}\sum_{i=1}^n \frac{\alpha_i^2}{\sigma_{\varepsilon i}^2}$, *and* $\bar{u}_t$ *and* $\bar{p}_t$ *are defined as*

$$\bar{u}_t \equiv \frac{1}{n}\sum_{i=1}^n \frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i u_{it} \quad \textit{and} \quad \bar{p}_t \equiv \frac{1}{n}\sum_{i=1}^n \frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i p_{it}.$$

This result indicates that with heterogeneity in the structure of demand shocks, there always exists an average price that fully reveals the information necessary to replicate the FCE dynamics. However, this is now a *weighted* average, where prices in sectors with more volatility idiosyncratic shocks or less sensitivity to the common component are given less weight. The intuition is straightforward: prices in more volatile or less sensitive sectors provide less informative signals about the common component, and so for the purpose of forecasting future values of this variable, those noisier signals need to be given less weight.

Proposition 9 raises the interesting possibility that there may exist a price index which reveals the right sufficient statistic needed to implement full communication plans, but firms instead observe a different price index. In this case, full revelation can fail, and the REE dynamics will differ from those in the FCE. To illustrate this possibility, Figure 1 shows the impulse responses of output for an economy with sectoral heterogeneity in which firms do not observe the appropriately weighted average price from Proposition 9. The economy has three sectors, which differ only in their sensitivity to the common component of demand, as in (16), with weights

$$(\alpha_1, \alpha_2, \alpha_3) = (-1, 1, 2).$$

Instead of observing its own demand shock u_{it} and the appropriately weighted average price $(\sum_{j=1}^{n} \alpha_j^2)^{-1} \sum_{i=1}^{n} \alpha_i p_{it}$, each firm instead observes its own demand shock u_{it} and the equally-weighted average price $n^{-1} \sum_{i=1}^{n} p_{it}$. In this example, the persistence of the common component is set to $\rho = 0.5$, and the values of all other parameters are the same as in Table 1 of Townsend (1983).

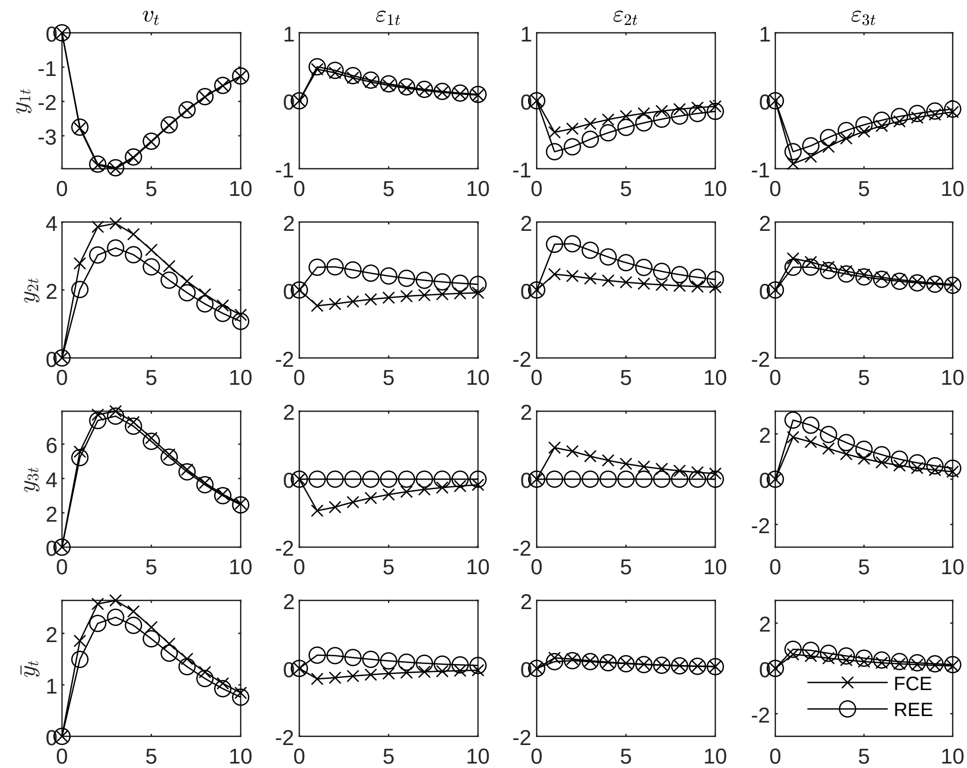

Figure 1: Consequences of observing the wrong price index in a 3-sector economy with $(\alpha_1, \alpha_2, \alpha_3) = (-1, 1, 2)$. The lines labeled REE show the responses of output when firms observe the equally-weighted average price instead of the appropriately weighted average price needed to support the FCE allocations. Parameter values: $\rho = 0.5$, $b_1 = 1$, $\beta = 0.96$, $f_0 = 0.2$, $f_1 = 0$, $f_2 = 0.8$, $\sigma_v^2 = \sigma_\varepsilon^2 = \sigma_\eta^2 = 1$.

Figure 1 shows that in the FCE a purely transitory idiosyncratic demand shock in sector 1 leads to a persistent decline in output in sector 2 (shown by the line marked with $\times$ in the (2,2) panel of the figure). Even under full communication, firms are not able to perfectly disentangle common and idiosyncratic shocks. The firm in sector

2 knows that demand in sector 1 has increased, but does not know whether this is because of a negative common shock or a positive idiosyncratic shock. The firm attaches some probability to the possibility that there was a negative common shock, with sectors 2 and 3 receiving offsetting positive idiosyncratic shocks, and so reduces production in response, with declining effects over time as the firm learns that the shock was not common.

When firms instead observe the equally-weighted average price, the firm in sector 2 still knows that its own demand has not changed, but now only observes an increase in the equally-weighted average price in response to the idiosyncratic shock in sector 1. It reasons that the increase in the average price could be due to a positive idiosyncratic shock in sector 1, but could also be due to a positive common shock together with an offsetting negative idiosyncratic shock in sector 2. This is because sector 3 is more sensitive (in absolute value) to the common component than sector 1, so the net effect of a common demand shock would be positive, also leading to an increase in the equally-weighted average price. In this numerical example, on balance the firm in sector 2 attaches greater probability to the possibility that the price increase is due to a positive common demand shock, and therefore responds by increasing rather than decreasing production.

This example illustrates how information revelation can fail to obtain simply as a result of sectoral heterogeneity, when firms do not observe the appropriate price index. In this case, the observed price index only provides a noisy observation of the ideal price index, where the noise depends on the differences in weights. If firms observe $\tilde{p}_t = \sum_{i=1}^{n} w_i p_{it}$, for some arbitrary sequence of weights $\{w_i\}_{i=1}^{n}$, this can be written as $\tilde{p}_t = \bar{p}_t + \sigma_\eta \eta_t$, where $\bar{p}_t$ is the ideal index from Proposition 9, and the error term

$$\sigma_\eta \eta_t \equiv \sum_{i=1}^{n} \left(w_i - \frac{1}{n} \frac{\sigma_\varepsilon^2 \alpha_i}{\sigma_{\varepsilon i}^2} \right) p_{it}$$

acts as aggregate "noise" in the observation of the ideal price index. Therefore, the reason that information revelation fails to obtain in this case is conceptually similar to the reason that it fails to obtain when observations of endogenous variables are contaminated by exogenous noise, which is the situation analyzed in the next section.

More generally, with arbitrary types of heterogeneity in the structure of demand, it can no longer be guaranteed that there will exist a single common price index that is sufficient for all firms to implement their full communication plans. The reason is that each firm will need to compute its own sufficient statistic of the demand shocks in order to optimally forecast its own demand. To reveal this statistic, each firm will require observations of its own unique price index. However, if firms are able to observe the history of all prices in the economy, this intuition suggests that it will still be possible for them to implement their full communication plans.

4 Infinite state problem

This section proves that the equilibrium of the Townsend model does not have a finite state representation when firms observe average prices with error, and explains the problem with evidence to the contrary from the existing literature. It also provides a new numerical procedure for solving models of this type in the frequency domain.

In the context of models with dispersed information, the "infinite regress problem" refers to a situation in which the action of each agent depends not only on its own forecast of the exogenous state of the economy, but also on its forecast of the average forecasts of others, and on its forecast of the average forecast of that average forecast, and so on, ad infinitum.[10] The reason why this problem is interesting is because it has the potential to amplify and propagate existing structural disturbances, or open the door for purely expectational disturbances to affect equilibrium outcomes. However, the technical challenge this problem introduces is that it may cause the equilibrium dynamics to fail to have have a finite-dimensional state-space representation, making it infeasible to use standard Kalman filtering formulas to compute agents' expectations.

In the case of full revelation, considered in the previous section, the infinite regress problem does not lead to an infinite state problem.[11] As discussed in the previous section, when firms observe an appropriately weighted price index, the REE dynamics of sectoral output, capital, and prices admit the same finite-state representation as in the FCE. However, for the same reason, the infinite regress problem does not play any role in affecting the equilibrium dynamics. Even though firms are required to form beliefs about demand shocks indirectly through observations of endogenous variables, the equilibrium dynamics are identical to what they would be in an economy where

[10]This differs from the infinite regress problem in bounded rationality contexts; cf. Conlisk (1996).

[11]Another situation in which the infinite regress problem does not lead to an infinite state problem is when agents do not learn from endogenous variables, as in Woodford (2003).

firms are able to observe demand shocks directly.

Based on revelation results of this type, the literature appears to have concluded that, for better or worse, the "new and exciting dynamics" envisioned by Townsend (1983) fail to appear in his model. Pearlman and Sargent (2005) summarize this view:

> "Townsend created [his] environment as a laboratory in which to study the effects of unleashing 'higher order beliefs.' He wanted to put [agents] into a setting in which they would have to estimate the beliefs of others in order to solve their own optimization and forecasting problems. The claim emerging from the string of papers just cited is that higher order beliefs disappear from this environment because there are so few sources of private information that prices can reveal all [agents'] private information. This result has both encouraging and discouraging aspects. Encouraging parts are that the equilibria of models like that of Townsend (1983), Section 8, are much easier to compute than Townsend originally thought, that standard recursive methods suffice to do the computations, [and] that the resulting equilibria have low-dimensional representations... A discouraging aspect is the fact that the dimension of the state-space is finite reflects the disappearance of the 'forecasting the forecasts of others problem' in equilibrium." (p.493)

However, this summary turns out to be misleading, for at least two reasons. The first is that the economy actually analyzed in Sec. 8 of Townsend (1983) is not one in which firms perfectly observe the appropriately weighted price index.[12] Instead, firms are assumed to observe it only *with error*. In terms of the baseline model described in Sec. 2 above, Townsend takes $n \to \infty$ and assumes that, in addition to their own demand shock u_{it}, firms observe

$$\tilde{p}_t = \bar{p}_t + \sigma_\eta \eta_t, \tag{17}$$

where $\bar{p}_t \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} p_{it}$, and the random variable η_t is jointly Gaussian but uncorrelated with the other disturbances in the model, uncorrelated over time, with mean zero and unit variance. In this case, the information revelation results from the previous section no longer apply (including the special case of Proposition 2 discussed

[12] Angeletos and Lian (2016) also point this out in footnote bk on p.1155.

by Pearlman and Sargent), and cannot be used to determine whether the infinite state problem discussed by Townsend appears under his own informational assumptions.

The second reason that the summary above is misleading is because papers that do follow Townsend in assuming that observations of average prices are contaminated with error do not prove any of the results described there. Two papers that may appear to suggest the opposite are Sargent (1991) and Kasa (2000). Sargent (1991) claims that the inclusion of "moving average components in agents' perceptions and of lagged innovations to agents' information in the state vector...enables [him] to formulate the equilibrium as a fixed point of a finite-dimensional operator" (pp.246-247). However, only numerical evidence is provided to support this claim, and it turns out to be incorrect (as Proposition 10 will show). Kasa (2000) presents the closed-form solution to the model, in Proposition 2.2.3, and it has a finite-dimensional state-space representation. However, the closed-form solution presented there is incorrect, essentially for the same reason discussed in Sec. 3 above: the procedure used to derive the solution fails to take into account that some observables can become informationally redundant in equilibrium.

The following result clarifies the situation, by proving that the infinite regress problem does indeed lead to an infinite state problem in the Townsend model. Similar results have been asserted in simpler settings, such as Chari (1979), Makarov and Rytchkov (2012), and Huo and Takayama (2023). Establishing a similar result in the Townsend model is more difficult due to its more complex dynamics, both in firms' observation process and in the underlying model structure.

Proposition 10. *Consider the economy from Sec. 2, with $n \to \infty$ and $s_{it} = (u_{it}, \tilde{p}_t)$ for all i and t, where $\tilde{p}_t$ is given by* (17)*. There does not exist a symmetric REE in which $\{y_{it}, k_{it}, p_{it}\}$ have finite-order ARMA representations.*

The structure of this proof is very similar to the proofs of Propositions 3 and 7, and basically amounts to a brute-force application of the infinite-dimensional method of undetermined coefficients described in Townsend (1983). The first step is to write the equilibrium law of motion of the endogenous signal as

$$\tilde{p}_t = A(L)v_t + B(L)\eta_t, \tag{18}$$

where $A(L)$ and $B(L)$ are one-sided operators, and use this to find the Wold representation of the observation vector $s_{it} = (u_{it}, \tilde{p}_t)$. Here, $\tilde{p}_t$ only depends on aggregate shocks due to the assumption of symmetry and the fact that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \varepsilon_{it} = 0$. The second step is to use the classical filtering formulas and structural equations of the model to compute the equilibrium fixed point equation

$$(A(L), B(L)) = T[(A(L), B(L))] \tag{19}$$

in closed form. The third step is to suppose to the contrary that $A(L)$ and $B(L)$ are rational functions of L, meaning that they can be written as ratios of polynomials with no common zeros, and use this hypothesis to rewrite the fixed point equation (19) in terms of those polynomials. The fourth step is to derive a contradiction, proving no operators $(A(L), B(L))$ satisfying (19) can be rational functions of L. This implies that neither $\{\tilde{p}_t\}$ nor the processes $\{y_{it}, k_{it}, p_{it}\}$ which depend on it can be expressed as finite-order ARMA processes.

4.1 Numerical procedure

In response to the infinite state problem, the literature has taken one of three different approaches to compute the solution of the model numerically. The first is to modify the information structure of agents in the model by assuming that all exogenous disturbances become common knowledge after a finitely many periods. This is the approach taken by Townsend (1983), originally proposed by Chari (1979). The second is to include only a finite number of higher-order expectations, as in Nimark (2017). The third is to use ARMA processes to numerically approximate the equilibrium dynamics, even though the endogenous variables do not have ARMA representations. This is the approach taken by Sargent (1991), and further developed by Han et al. (2022), Adams (2022), and Huo and Takayama (2023).

As a complementary alternative to these, this paper proposes a new numerical procedure that does not rely on ARMA approximations.[13] The main conceptual advantage of our approach is that it is a direct implementation of the same procedure that is used to prove the our theoretical results. The basic idea is to iterate on the equilibrium fixed point equation of the model in the frequency domain rather than in the time domain. The difference relative to Han et al. (2022) is that the classical Wiener-Kolmogorov filter is used to compute forecasts instead of the Kalman fitler, as is done theoretically in the proofs of Propositions 3, 7, and 10.

[13] A parallel procedure for models with rationally inattentive agents is presented in Jurado (2023).

The procedure involves iterating on the equilibrium fixed point equation of the model. To derive that equation, begin by writing firms' perceived law of motion for the noisy endogenous price signal in any symmetric REE as

$$\tilde{p}_t = A(L)v_t + B(L)\eta_t, \tag{20}$$

where $A(L)$ and $B(L)$ are one-sided operators. In terms of these operators, the law of motion of the observation vector $s_{it} = (u_{it}, \tilde{p}_t)$ is given by

$$s_{it} = \begin{bmatrix} H_v(L) & 0 & H_\varepsilon(L) \\ A(L) & B(L) & 0 \end{bmatrix} \begin{bmatrix} v_t \\ \eta_t \\ \varepsilon_{it} \end{bmatrix} \equiv M(L)e_{it}, \tag{21}$$

where $e_{it} \equiv (v_t, \eta_t, \varepsilon_{it})$. To describe the procedure, the exogenous laws of motion in (4) and (18) have been generalized to $u_{it} = H_v(L)v_t + H_\varepsilon(L)\varepsilon_{it}$ and $\tilde{p}_t = \bar{p}_t + H_\eta(L)\eta_t$, where $H_v(L)$, $H_\varepsilon(L)$, and $H_\eta(L)$ are one-sided and invertible into the past.

Letting $s_{it} = \Gamma(L)w_{it}$ denote the Wold representation associated with the law of motion (21), the Hansen and Sargent (1981) formula implies that

$$\sum_{j=1}^{\infty}(\beta\lambda)^j E_{it}u_{i,t+j} = \frac{\beta\lambda}{L-\beta\lambda}\begin{bmatrix} 1 & 0 \end{bmatrix}(\Gamma(L) - \Gamma(\beta\lambda))\Gamma(L)^{-1}s_{it}.$$

Substituting this into the policy function (6) and the demand curve (2), and then averaging across sectors delivers the implied actual law of motion

$$\tilde{p}_t = H_v(L)v_t + H_\eta(L)\eta_t - \frac{\beta\lambda}{L-\beta\lambda}\begin{bmatrix} 1 & 0 \end{bmatrix}(\Gamma(L) - \Gamma(\beta\lambda))\Gamma(L)^{-1}\bar{s}_t. \tag{22}$$

Matching coefficients in the perceived and actual laws of motion (20) and (22) delivers the equilibrium fixed point equation

$$\begin{aligned}\begin{bmatrix} A(L) & B(L) \end{bmatrix} &= \begin{bmatrix} H_v(L) & H_\eta(L) \end{bmatrix} \\ &\quad - \frac{b_1 f_0^2 \beta\lambda L}{f_2(1-\lambda L)(L-\beta\lambda)}\begin{bmatrix} 1 & 0 \end{bmatrix}(\Gamma(L) - \Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} H_v(L) & 0 \\ A(L) & B(L) \end{bmatrix}.\end{aligned} \tag{23}$$

The numerical procedure involves iterating on this equation by representing the operators $A(L)$ and $B(L)$ in the frequency domain. This means viewing operators of the form $A(L) = \sum_{s=0}^{\infty} A_s L^s$ as functions of the real parameter $\omega \in [-\pi, \pi]$ by defining $a(\omega) \equiv \lim_{r\to 1} A(re^{-\mathrm{i}\omega})$, where i is the imaginary unit and ω represents the "frequency." (The convention here is that lower-case letters denote functions of ω and upper-case letters denote functions of L.) Numerically, $a(\omega)$ can be represented on a discrete grid of frequencies $\omega_1, \omega_2, \ldots, \omega_N$ by the sequence $\{a(\omega_1), a(\omega_2), \ldots, a(\omega_N)\}$. Using approximations of this type, the algorithm can be described as follows.

Algorithm 1. Initialize the functions $(a^{(n)}(\omega), b^{(n)}(\omega))$ on a discrete grid over $[-\pi, \pi]$.

(1) Substitute $(a^{(n)}(\omega), b^{(n)}(\omega))$ into (21) to compute $m^{(n)}(\omega)$.

(2) Use the factorization procedure of Tunnicliffe-Wilson (1972) to compute $\gamma^{(n)}(\omega)$.

(3) Use (23) to compute the updated functions $(a^{(n+1)}(\omega), b^{(n+1)}(\omega))$.

(4) Repeat (1)-(3) until $\|(a^{(n+1)}(\omega), b^{(n+1)}(\omega)) - (a^{(n)}(\omega), b^{(n)}(\omega))\|$ is acceptably low.

Once numerical approximations of the functions $a(\omega)$ and $b(\omega)$ have been obtained by means of this algorithm, the coefficient sequences $\{A_s\}$ and $\{B_s\}$, which represent the impulse responses of $\tilde{p}_t$ to the disturbances v_t and η_t, can be recovered using the inverse Fourier transform; i.e. $A_s = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\mathrm{i}\omega s} a(\omega) d\omega$. The Fast Fourier Transform algorithm provides a numerically efficient method of approximating the Fourier coefficients of a square-integrable function. Implementations of this algorithm are available in most numerical programming packages. In Matlab, this algorithm is implemented by the built-in function `ifft`, which accepts both univariate and multivariate inputs.

In Step 2 of Algorithm 1, the factorization procedure of Tunnicliffe-Wilson (1972) takes the place usually occupied by the Kalman filter in finding the Wold factor associated with firms' observation process. This procedure computes $\gamma(\omega) = \Gamma(e^{-\mathrm{i}\omega})$ by directly factorizing the spectral density $f(\omega) \equiv M(e^{-\mathrm{i}\omega})M(e^{-\mathrm{i}\omega})^*$ in the frequency domain, using a matrix version of Newton's method for obtaining square roots. Although we save a full comparison of numerical properties of different algorithms for future work, the Online Appendix presents evidence that our algorithm offers some advantages relative to one proposed by Han et al. (2022), at least in the context of this model.

The next subsection uses Algorithm 1 to explore how the presence of endogenous signals affects the equilibrium dynamics of the model.

4.2 Effects of endogenous signals

The key economic mechanism in the Townsend model is that agents learn about the underlying state of the economy through imperfect observations of aggregate variables, which act as endogenous signals. But how much of an effect does this mechanism have on the equilibrium dynamics? Proposition 10 implies that one effect is that the dynamics cannot be represented by a finite-dimensional system. However, it is not clear from this theoretical result whether this difference is either quantitatively or economically substantial. The purpose of this subsection is to present results from a simple numerical exercise that helps address this question.

The exercise is to compare the equilibrium dynamics in the Townsend model with endogenous signals to alternative versions of the model without them. For this purpose, two alternative versions of the model serve as relevant benchmarks. The first is one in which firms have full information about all underlying disturbances. In this version of the model, there is no learning from endogenous variables because there is no learning at all. Firms face no uncertainty about the current state of the economy, and all information is common. The second is one in which firms do face uncertainty about the current state of the economy, but they receive only exogenous signals about it, as in Woodford (2003). In this version, there is learning, but not from endogenous variables.

More specifically, consider three versions of the Townsend economy, where the observation vector is specified variously as

(1) Full information: $s_{it} = (\varepsilon_t, \theta_t, \eta_t)$, where $\varepsilon_t \equiv (\varepsilon_{1t}, \varepsilon_{2t}, \dots)$.

(2) Exogenous signal: $s_{it} = (u_{it}, \tilde{\theta}_t)$, where $\tilde{\theta}_t = \theta_t + \sigma_\eta \eta_t$.

(3) Endogenous signal: $s_{it} = (u_{it}, \tilde{p}_t)$, where $\tilde{p}_t = \bar{p}_t + \sigma_\eta \eta_t$.

Figure 2 plots the impulse responses of the average level of output, the average price, and the average error in estimating θ_t to the aggregate disturbances in the model in each of these three versions of the model. The parameter values are the ones from Table 1 of Townsend (1983). The relevant comparison is in the difference between the responses under each of the different informational assumptions.

What Figure 2 shows is that there is a much larger difference in dynamics between the full information economy and the two dispersed information economies, than between the two dispersed information economies themselves. Indeed, the responses

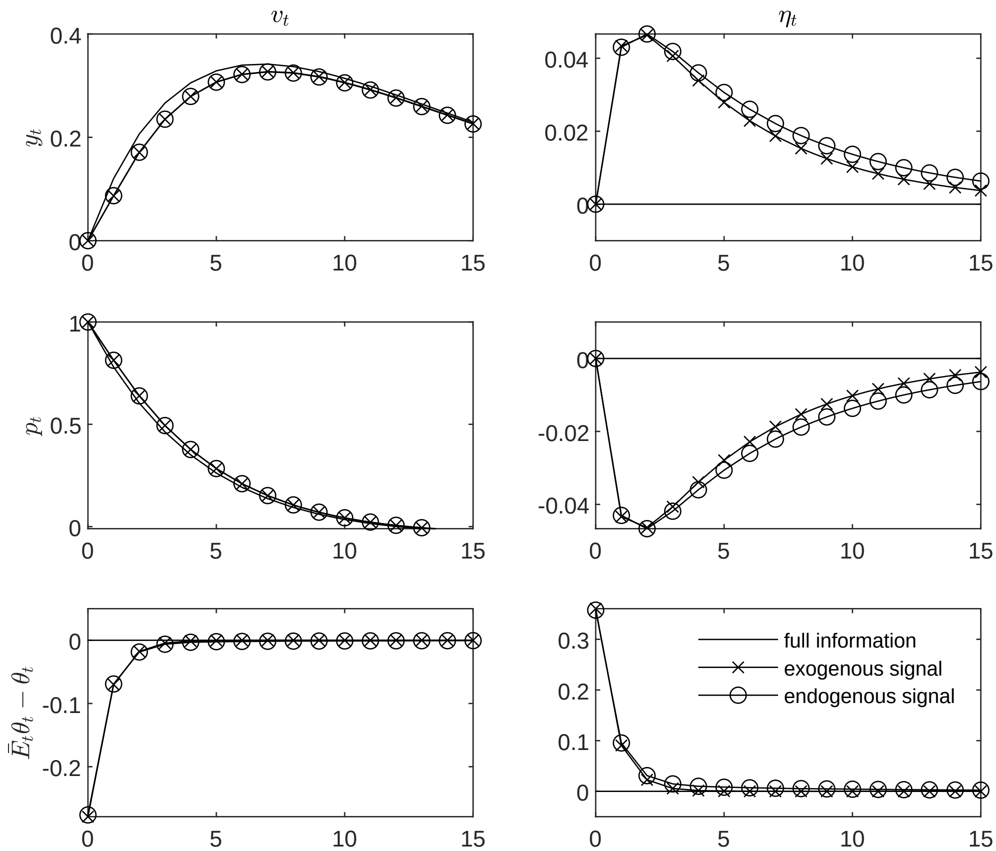

Figure 2: Effects of endogenous signals on aggregates. This figure shows the responses of average level of output, prices, and the average estimation error of θ_t in response to the common demand disturbance v_t and the signal noise disturbance η_t. Parameter values: $\rho = 0.9$, $b_1 = 1$, $\beta = 0.96$, $f_0 = 0.2$, $f_1 = 0$, $f_2 = 0.8$, $\sigma_v^2 = \sigma_\varepsilon^2 = \sigma_\eta^2 = 1$.

in the exogenous signal economy are very similar to those in the endogenous signal economy, despite the fact that the dynamics admit a finite-dimensional representation in the first case but not the second. Nevertheless, the endogenous signal economy does exhibit the most persistence, both with respect to the aggregate demand shock (though visually imperceptible in the figure) and the aggregate noise shock.[14]

The similarity between the two dispersed information models can also be seen in their implications for the dynamics of higher-order expectations. Figure 3 illustrates the implied dynamics of higher-order expectations of θ_t in response to both aggregate

[14]In contrast to the discussion of Kasa (2000), the bottom two panels show that the model does exhibit waves of optimism and pessimism that last for more than one period, consistent with Figures 4 and 5 of Townsend (1983).

shocks. Defining

$$\bar{E}_t^{(0)}\theta_t \equiv \theta_t \quad \text{and} \quad \bar{E}_t^{(k)}\theta_t \equiv \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} E_{it}[\bar{E}_t^{(k-1)}\theta_t],$$

the figure plots the response of $\bar{E}_t^{(k)}\theta_t$ for various values of the parameter k. All panels show that in response to the shocks, higher order expectations converge more slowly towards the true response (solid line). In response to noise shocks, this means that higher order expectations increase by more on impact, and more slowly adjust to zero. Comparing the top and bottom rows of panels, it can be seen that the degree of sluggishness in the higher-order expectations is moderately greater in the endogenous signal model, but overall the dynamics are quite similar.

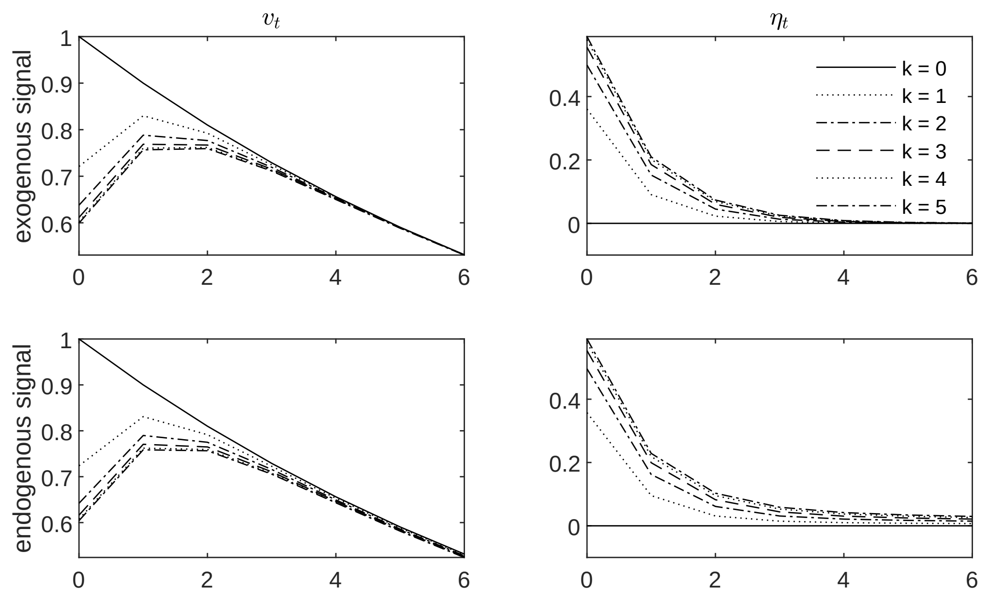

Figure 3: Effects of endogenous signals on higher-order expectations. This figure shows the responses of for higher-order expectations of the common demand component, $\bar{E}_t^{(k)}\theta_t$, for various values of k. The case $k = 0$ indicates the response of θ_t itself. Parameter values are the same as in Figure 2.

Even though the exogenous signal economy has very similar predictions in terms of the time-series dynamics of model variables, there are at least three important caveats. The first is a reminder that this conclusion is specific both to the particular model and parameter values chosen. In terms of the model parameters, some experimentation with different values suggests that it is difficult to make the responses in the two dispersed information economies differ by much more than they do in the figure. In terms of the model itself, the focus here is narrowly on the Townsend model, and while this has been a helpful laboratory in the dispersed information literature for some time, obviously nothing rules out the possibility that the effect of endogenous signals may be larger in other environments. The exercise in this section does suggest, however, that a numerical comparison of this type would be helpful for isolating the contribution of the endogenous learning mechanism in other environments.

The second caveat is that even if the exogenous and endogenous signal economies have similar dynamics, the latter economy imposes greater *discipline* on agents' information structures. Supposing that firms observe an exogenous signal $\tilde{\theta}_t = C(L)v_t + D(L)\eta_t$, it is not obvious what form the operators $C(L)$ and $D(L)$ should take. Figure 2 chooses $C(L) = \sigma_v/(1 - \rho L)$ and $D(L) = \sigma_\eta$, and while this choice may seem natural from a statistical modeling perspective, there is no economic justification for it. In the endogenous signal economy, by contrast, the mapping from the average price to the shocks is endogenously determined by other model assumptions. Moreover, it is not difficult to see that for any endogenous signal economy, there exists an observationally equivalent exogenous signal economy: simply set $C(L)$ and $D(L)$ equal to the equilibrium values of $A(L)$ and $B(L)$ implied by the endogenous signal economy. While it may be surprising that the *particular* forms of $C(L)$ and $D(L)$ chosen in Figure 2 happen to replicate the dynamics of the endogenous information economy fairly well, the fact that there exist *some* operators that do this is not.

The third caveat is that exogenous signal economies do not permit analysis of how changes in model structure, including policy, affect agents' information. This is a central aspect of the macroeconomic literature on endogenous information choice, especially the literature on rational inattention following Sims (2003). While agents in the Townsend model do not choose their information sets optimally subject to information processing constraints, the fact that they are still required to learn from endogenous variables does mean that their information sets endogenously respond to structural changes, unlike in an exogenous signal economy.

5 Conclusion

Prices are often referred to as signals. But in most modern macroeconomic models, they play no formal role in transmitting information. One of the first dynamic models that explicitly considers this mechanism is the one developed by Townsend (1983).

However, the subtle technical and conceptual issues that this model raises have led to a degree of confusion in the subsequent literature. This paper has revisited this influential model to help provide some precision and clarity.

On the one hand, existing results about information revelation in this model are not stated as strongly as they could be. A single price index can reveal a substantial amount of information, fully revealing all essential information even in the presence of a great deal of heterogeneity, as shown in Propositions 2, 6, 8, and 9. On the other hand, existing results about information revelation in this model are stated more strongly than they should be. Realistic types of heterogeneity or noise in the observation of prices can prevent full revelation, and can lead to a situation in which the equilibrium state vector can become infinite dimensional, as in Proposition 10.

From a methodological perspective, the proofs provided in this paper can be read as a step-by-step guide for how to prove similar results in other models, and the numerical procedure described in Section 4.1 is broadly applicable. Hopefully these contributions will help reduce barriers to entry for working on models with endogenous signals, especially as new evidence on higher-order expectations, such as the survey data from Coibion et al. (2021), makes it possible to directly discipline models of this type in ways that were not feasible when they were originally formulated.

References

Abel, A. B. and O. J. Blanchard (1983). An intertemporal model of saving and investment. *Econometrica 51*(3), 675–692.

Acharya, S., J. Benhabib, and Z. Huo (2021). The anatomy of sentiment-driven fluctuations. *Journal of Economic Theory 195*, 105280.

Adams, J. J. (2022). Macroeconomic models with incomplete information and endogenous signals. University of Florida Working Paper.

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

Angeletos, G.-M. and C. Lian (2016). Chapter 14 - incomplete information in macroeconomics: Accommodating frictions in coordination. Volume 2 of *Handbook of Macroeconomics*, pp. 1065–1240. Elsevier.

Bacchetta, P. and E. Van Wincoop (2006). Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review 96*(3), 552–576.

Bernhardt, D., P. Seiler, and B. Taub (2010). Speculative dynamics. *Economic Theory 44*(1), 1–52.

Chari, V. V. (1979). Linear Stochastic Models with Decentralized Information. Carnegie-Melon Discussion Paper No. 458.

Coibion, O., Y. Gorodnichenko, S. Kumar, and J. Ryngaert (2021). Do You Know that I Know that You Know...? Higher-Order Beliefs in Survey Data. *The Quarterly Journal of Economics 136*(3), 1387–1446.

Conlisk, J. (1996). Why Bounded Rationality? *Journal of Economic Literature 34*(2), 669–700.

Futia, C. A. (1981). Rational Expectations in Stationary Linear Models. *Econometrica 49*(1), 171–192.

Grossman, S. (1976). On the efficiency of competitive stock markets where trades have diverse information. *The Journal of Finance 31*(2), 573–585.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American Economic Review 70*(3), 393–408.

Han, Z., F. Tan, and J. Wu (2022). Analytic policy function iteration. *Journal of Economic Theory 200*, 105395.

Hansen, L. P. and T. J. Sargent (1981). A note on wiener-kolmogorov prediction formulas for rational expectations models. *Economics Letters 8*(3), 255–260.

He, H. and J. Wang (1995). Differential information and dynamic behavior of stock trading volume. *The Review of Financial Studies 8*(4), 919 – 972.

Hellwig, C. and V. Venkateswaran (2009). Setting the right prices for the wrong reasons. *Journal of Monetary Economics 56*, S57–S77.

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of Economic Theory 22*(3), 477–498.

Huo, Z. and M. Pedroni (2023). Dynamic information aggregation: Learning from the past. *Journal of Monetary Economics 136*, 107–124.

Huo, Z. and N. Takayama (2023). Rational Expectations Models with Higher Order Beliefs. Yale Univeristy Working paper.

Jurado, K. (2023). Rational inattention in the frequency domain. *Journal of Economic Theory 208*, 105604.

Kasa, K. (2000). Forecasting the forecasts of others in the frequency domain. *Review of Economic Dynamics 3*(4), 726 – 756.

Kasa, K., T. B. Walker, and C. H. Whiteman (2014). Heterogeneous Beliefs and Tests of Present Value Models. *The Review of Economic Studies 81*(3), 1137–1163.

Kimbrough, K. P. (1984). Aggregate information and the role of monetary policy in an open economy. *Journal of Political Economy 92*(2), 268–285.

King, R. G. (1982). Monetary policy and the information content of prices. *Journal of Political Economy 90*(2), 247 – 279.

Kreps, D. M. (1977). A note on "fulfilled expectations" equilibria. *Journal of Economic Theory 14*(1), 32–43.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review 99*(5), 2050–84.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory 4*(2), 103–124.

Lucas, R. E. and E. C. Prescott (1971). Investment under uncertainty. *Econometrica 39*(5), 659–681.

Makarov, I. and O. Rytchkov (2012). Forecasting the forecasts of others: Implications for asset pricing. *Journal of Economic Theory 147*(3), 941–966.

Marcet, A. and T. J. Sargent (1989). Convergence of least-squares learning in environments with hidden state variables and private information. *Journal of Political Economy 97*(6), 1306–1322.

Melosi, L. (2016). Signalling Effects of Monetary Policy. *The Review of Economic Studies 84*(2), 853–884.

Miao, J., J. Wu, and E. R. Young (2021). Macro-financial volatility under dispersed information. *Theoretical Economics 16*(1), 275–315.

Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Nimark, K. P. (2014). Man-bites-dog business cycles. *American Economic Review 104*(8), 2320–67.

Nimark, K. P. (2017). Dynamic higher order expectations. Cornell University Working Paper.

Pearlman, J., D. Currie, and P. Levine (1986). Rational expectations models with partial information. *Economic Modelling 3*(2), 90–105.

Pearlman, J. G. and T. J. Sargent (2005). Knowing the forecasts of others. *Review of Economic Dynamics 8*(2), 480–497. Monetary Policy and Learning.

Radner, R. (1979). Rational expectations equilibrium: Generic existence and the information revealed by prices. *Econometrica 47*(3), 655–678.

Rondina, G. and T. Walker (2023). Equivalence results with endogenous signal extraction. CAEPR Working Papers 2023-007, Center for Applied Economics and Policy Research, Department of Economics, Indiana University Bloomington.

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

Rozanov, Y. A. (1967). *Stationary Random Processes*. San Francisco, CA: Holden-Day.

Sargent, T. J. (1991). Equilibrium with signal extraction from endogenous variables. *Journal of Economic Dynamics and Control 15*(2), 245–273.

Sims, C. A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics 50*(3), 665–690.

Singleton, K. J. (1987). Asset prices in a time-series model with disparately informed, competitive traders. In W. A. Barnett and K. J. Singleton (Eds.), *New Approach to Monetary Economics*, pp. 2867–2872. Proceedings of the Second International Symposium in Economic Theory and Econometrics.

Taub, B. (1989). Aggregate fluctuations as an information transmission mechanism. *Journal of Economic Dynamics and Control 13*(1), 113–150.

Townsend, R. M. (1983). Forecasting the forecasts of others. *Journal of Political Economy 91*(4), 546–588.

Tunnicliffe-Wilson, G. (1972). The factorization of matricial spectral densities. *SIAM Journal on Applied Mathematics 23*(4), 420–426.

Whittle, P. (1983). *Prediction and Regulation by Linear Least-Square Methods.* Minneapolis, MN: University of Minnesota Press.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. In P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (Eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, pp. 25–58. Princeton, NJ: Princeton University Press.

Woodford, M. (2013). Macroeconomic analysis without the rational expectations hypothesis. *Annual Review of Economics 5*(1), 303–346.

A Proofs

Proof of Proposition 1. Substituting (2) into (1) implies

$$E_{it}[(1-\lambda_1 L)(1-\lambda_2 L)k_{i,t+2}] = -\frac{f_0}{f_2} E_{it} u_{i,t+1}, \tag{24}$$

where λ_1 and λ_2 are zeros of $\mathcal{P}(\lambda) \equiv \lambda^2 - (1+\beta^{-1}+(b_1 f_0^2 + f_1)/f_2)\lambda + \beta^{-1}$. They satisfy $0 < \lambda_1 < 1 < \lambda_2$, because $\mathcal{P}(0) > 0$, $\mathcal{P}(1) < 0$, and $\mathcal{P}(\lambda) > 0$ for large λ. The unique stationary solution to (24) is (6), where $\lambda \equiv \lambda_1$. Since $s_{it} = u_t$ is exogenous, $E_{it} u_{i,t+j}$ exists and is unique for all j by the orthogonal projection theorem. By (2),

$$E_{it} u_{i,t+j} = \rho^j E_{it} \theta_t. \tag{25}$$

The vector $u_t = (u_{1t}, \dots, u_{nt})$ contains the same information as $(\bar{u}_t, \hat{u}_{1t}, \dots, \hat{u}_{2t})$, where $\hat{u}_{it} \equiv u_{it} - \bar{u}_t$, since the two are related by a non-degenerate linear transformation. Since the processes $\{\hat{u}_{1t}, \dots, \hat{u}_{nt}\}$ are independent of $\{\theta_t\}$,

$$E_{it}\theta_t = E(\theta_t | u^t) = E(\theta_t | \bar{u}^t, \hat{u}_1^t, \dots, \hat{u}_n^t) = E(\theta_t | \bar{u}^t). \tag{26}$$

By (4), $\bar{u}_t = \theta_t + \sigma_\varepsilon \bar{\varepsilon}_t$, where $\bar{\varepsilon}_t \equiv \frac{1}{n}\sum_{i=1}^n \varepsilon_t$ is white noise with variance $1/n$. The Wold factor associated with the spectral density of $\{\bar{u}_t\}$ is $H(L) = \sqrt{\frac{\rho\sigma_\varepsilon^2}{\phi n}\frac{(1-\phi L)}{(1-\rho L)}}$, where ϕ is defined in the proposition. By the Wiener-Kolmogorov filtering equation,

$$E(\theta_t | \bar{u}^t) = \left[\frac{\sigma_v^2}{(1-\rho L)(1-\rho L^{-1})} H(L^{-1})^{-1}\right]_+ H(L)^{-1}\bar{u}_t = \frac{(1-\phi/\rho)}{(1-\phi L)}\bar{u}_t, \tag{27}$$

where $[\,\cdot\,]_+$ projects onto the space spanned by non-negative powers of L, and the second equality uses the fact that $\sigma_v^2 = \frac{\rho\sigma_\varepsilon^2}{\phi n}(1-\phi\rho)(1-\phi/\rho)$, by definition of ϕ. Substituting (25), (26), and (27) into (6) delivers the final expression for $k_{i,t+1}$.

Proof of Proposition 2. It is sufficient to verify that the operator on the right side of (7) is invertible. This is true because $\mathcal{P}(\mu) \equiv \mu^2 - (\lambda + \phi + b_1 f_0 \omega)\mu + \lambda\phi$ has two inside zeros, since $\mathcal{P}(0) > 0$, $\mathcal{P}(\lambda) < 0$, and $\mathcal{P}(1) = (1-\lambda)(1-\phi) - b_1 f_0 \omega > 0$.[15] The last inequality follows because, by definition of λ and ϕ,

$$(1-\lambda)(1-\phi) = \frac{b_1 f_0^2 + f_1}{f_2}\frac{\beta\lambda}{1-\beta\lambda}\frac{\sigma^2 + (1-\rho)\sigma_\varepsilon^2}{\sigma^2+\sigma_\varepsilon^2} > \frac{b_1 f_0^2}{f_2}\frac{\beta\lambda\rho}{1-\beta\lambda\rho}\frac{\sigma^2}{\sigma^2+\sigma_\varepsilon^2} = b_1 f_0 \omega.$$

Proof of Proposition 3. The proof has four steps.

Step 1. *Prove that $\bar{p}_t = A(L)\bar{u}_t$ in any symmetric REE, with $A(L)$ one-sided.* In any symmetric REE, $k_{i,t+1}$ is measurable with respect to $s_i^t = (u_i^t, \bar{p}^t)$, so $k_{i,t+1} = B_u(L)u_{it} + B_p(L)\bar{p}_t$ for some one-sided operators $B_u(L)$ and $B_p(L)$, which don't depend on i by symmetry. Substituting this expression for $k_{i,t+1}$ into (2), averaging across i, and solving for $\bar{p}_t$ implies $\bar{p}_t = \frac{1-b_1 f_0 B_u(L)L}{1+b_1 f_0 B_p(L)L}\bar{u}_t \equiv A(L)\bar{u}_t$. And, since $\bar{p}_t$ cannot be correlated with future disturbances, $A(L)$ must be one-sided.

Step 2. *Find the Wold representation of the observation vector $s_{it} = \Gamma(L)w_{it}$.*

[15] "Inside zero" is shorthand for "zero inside the unit circle;" the same goes for "outside zero."

The law of motion for s_{it} is

$$s_{it} = \frac{1}{1-\rho L}\begin{bmatrix} \sigma_v & \sigma_\varepsilon(1-\rho L)\iota_i' \\ \sigma_v A(L) & \sigma_\varepsilon(1-\rho L)A(L)\frac{1}{n}1_n' \end{bmatrix}\begin{bmatrix} v_t \\ \varepsilon_t \end{bmatrix} \equiv \frac{1}{1-\rho L}M_i(L)e_t,$$

where ι_i is a vector of zeros with a one in the i-th position, and 1_n is an n-dimensional vector of ones. Given $M_i(L)$, the Wold factor $\Gamma(L)$ can be computed using the procedure from pp.44-47 of Rozanov (1967). The result is

$$\Gamma(L) = \frac{(1+\vartheta^2)^{-1/2}}{(1-\rho L)} \tag{28}$$

$$\times \begin{bmatrix} \vartheta\sqrt{\frac{\rho}{\alpha}}\sigma_\varepsilon(L-\alpha) & -\sqrt{\frac{\rho}{\alpha}}\sigma_\varepsilon(1-\alpha L) \\ \vartheta\sqrt{\frac{\rho}{\alpha}}\frac{\alpha}{\phi}\frac{\sigma_\varepsilon}{n}\frac{(1-\phi L)(L-\phi)}{(1-\alpha L)}A(L)+\gamma(L) & -\sqrt{\frac{\rho}{\alpha}}\frac{\alpha}{\phi}\frac{\sigma_\varepsilon}{n}\frac{(1-\phi L)(L-\phi)}{(L-\alpha)}A(L)+\vartheta\gamma(L)\frac{1-\alpha L}{L-\alpha} \end{bmatrix},$$

where

$$\vartheta \equiv \sqrt{\frac{\rho}{\alpha}\frac{\alpha}{\phi}\frac{\sigma_\varepsilon}{n}}\frac{(1-\alpha\phi)(\alpha-\phi)}{(1-\alpha^2)}\frac{A(\alpha)}{\gamma(\alpha)}, \tag{29}$$

$\gamma(L)$ is the univariate Wold factor that satisfies

$$\gamma(L)\gamma(L^{-1}) = A(L)A(L^{-1})\frac{\alpha}{\phi}\frac{\sigma_\varepsilon^2}{n}\left(1-\frac{1}{n}\right)\frac{(1-\phi L)(1-\phi L^{-1})(1-\rho L)(1-\rho L^{-1})}{(1-\alpha L)(1-\alpha L^{-1})}, \tag{30}$$

and α and ϕ solve the quadratic equations

$$\rho\sigma_\varepsilon^2\alpha^2 - (\sigma_\varepsilon^2(1+\rho^2)+\sigma_v^2)\alpha + \rho\sigma_\varepsilon^2 = 0, \quad \rho\sigma_\varepsilon^2\phi^2 - (\sigma_\varepsilon^2(1+\rho^2)+n\sigma_v^2)\phi + \rho\sigma_\varepsilon^2 = 0, \tag{31}$$

respectively, and satisfy the inequalities $0 < \phi < \alpha < \rho$. It is possible to verify that the roots of the determinant of $\Gamma(L)$ in (28) lie inside the unit circle (so $\Gamma(L)$ is invertible) and $\Gamma(L)\Gamma(L^{-1})' = \frac{1}{(1-\rho L)(1-\rho L^{-1})}M(L)M(L^{-1})'$.

Step 3. *Find the equilibrium fixed point equation* $A(L) = T[A(L)]$. Using the Hansen and Sargent (1981) formula and averaging across sectors,

$$\sum_{j=1}^{\infty}(\beta\lambda)^j\bar{E}_t u_{i,t+j} = \frac{\beta\lambda}{L-\beta\lambda}\left[\,1 \quad 0\,\right](\Gamma(L)-\Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} 1 \\ A(L) \end{bmatrix}\bar{u}_t.$$

Substituting this into (9) and matching coefficients on $\bar{u}_t$ implies

$$A(L) = 1 - \frac{b_1 f_0^2}{f_2} \frac{\beta\lambda L}{(1-\lambda L)(L-\beta\lambda)} \begin{bmatrix} 1 & 0 \end{bmatrix} (\Gamma(L) - \Gamma(\beta\lambda))\Gamma(L)^{-1} \begin{bmatrix} 1 \\ A(L) \end{bmatrix}.$$

Substituting in the expression for $\Gamma(L)$ in (28) and rearranging,

$$A(L) = \frac{(1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 L\psi(L)}{(1-\lambda L)(1-\alpha L)(L-\alpha) + b_0(1-\alpha^2)\vartheta\sigma_\varepsilon\sqrt{\frac{\alpha}{\rho}}\left(1-\frac{1}{n}\right)\frac{(1-\rho L)^2(L-\rho)L}{\gamma(L)(1-\alpha L)}}, \tag{32}$$

where $\psi(L) \equiv \vartheta^2(1-\alpha L)(1-\alpha\rho) + (L-\alpha)(\rho-\alpha)$, $b_0 \equiv \frac{b}{1+\vartheta^2}$, and $b \equiv \frac{b_1 f_0^2 \beta\lambda}{f_2(1-\beta\lambda\rho)} > 0$. By the definition of λ in Proposition 1,

$$b < 1 - \lambda. \tag{33}$$

Step 4. *Prove that $A(L)$ is invertible.* First, $\vartheta \neq 0$. If not, (32) implies

$$A(L) = \frac{(1-\lambda L)(1-\alpha L) - bL(\rho-\alpha)}{(1-\lambda L)(1-\alpha L)},$$

and (29) implies $A(\alpha) = 0$, so α must be a zero of the numerator on the right. But this is not possible, because (33) implies $b\alpha(\rho-\alpha) < (1-\alpha\lambda)(1-\alpha^2)$.

Second, it is helpful to rewrite (32) in terms of the Wold factor $H(L)$ that satisfies

$$H(L)H(L^{-1}) = A(L)A(L^{-1}). \tag{34}$$

By (30), $\gamma(L) = \sqrt{\frac{\sigma_\varepsilon^2}{n}\frac{\alpha}{\phi}\left(1-\frac{1}{n}\right)}\frac{(1-\phi L)(1-\rho L)}{(1-\alpha L)}H(L)$. Substituting this into (32) implies

$$A(L) = \frac{H(L)(1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 H(L)L\psi(L)}{H(L)(1-\lambda L)(1-\alpha L)(L-\alpha) + b_0(1-\alpha^2)\vartheta\sqrt{\frac{\phi(n-1)}{\rho}}\frac{(1-\rho L)(L-\rho)L}{(1-\phi L)}}. \tag{35}$$

Third, suppose $A(L)$ is not invertible, and has at least one inside zero (not zero, since $A(0) = 1$). By (34), $H(L)$ has at least one outside zero not shared by $A(L)$. This is possible only if the zero is $1/\rho$; otherwise the outside zero of $H(L)$ in the numerator of (35) will not cancel on the denominator, and will be a zero of $A(L)$. Therefore, $A(L)$ has a zero at $L = \rho$ of multiplicity one. By (34),

$$H(L) = \frac{1-\rho L}{L-\rho}A(L). \tag{36}$$

Substituting this into (35) and solving for $A(L)$,

$$A(L) = 1 - b_0 L \frac{(1-\phi L)\psi(L) + \vartheta(1-\alpha^2)\sqrt{\frac{\phi(n-1)}{\rho}}(L-\rho)^2}{(1-\phi L)(1-\lambda L)(1-\alpha L)(L-\alpha)}. \tag{37}$$

Equation (37) and the hypothesis that $A(\rho) = 0$ provide an expression for ϑ^2,

$$\vartheta^2 = -\frac{(\rho-\alpha)}{(1-\alpha\rho)} \frac{(1-\lambda\rho)(1-\alpha\rho) - b\rho(\rho-\alpha)}{(1-\lambda\rho)(\rho-\alpha) - b\rho(1-\alpha\rho)}, \tag{38}$$

where $b_0 = b/(1+\vartheta^2)$ has been used to substitute out b_0. Since $A(L)$ is one-sided, the numerator of the fraction on the right side of (37) must have a zero at $L = \alpha$ to cancel the factor $L - \alpha$ in the denominator. This provides a second expression for ϑ^2,

$$\vartheta^2 = \frac{(\alpha-\phi)(\rho-\alpha)^3}{(1-\alpha\phi)(1-\alpha\rho)^3}, \tag{39}$$

where it has been used that $\frac{\phi(n-1)}{\rho} = \frac{(1-\alpha\phi)(\alpha-\phi)}{(1-\alpha\rho)(\rho-\alpha)}$ by (31). Equating the two expressions for ϑ^2 in (38) and (39), and using (33) to eliminate b, implies

$$\frac{(\alpha-\phi)(\rho-\alpha)^2}{(1-\alpha\phi)(1-\alpha\rho)^2} > -\frac{(1-\alpha\rho) - \rho(\rho-\alpha)}{(\rho-\alpha) - \rho(1-\alpha\rho)} = \frac{1}{\alpha}.$$

Since $(\rho - \alpha) < (1 - \alpha\rho)$, this inequality implies $(\alpha - \phi)/(1 - \alpha\phi) > 1/\alpha$, which is a contradiction, because $\alpha(\alpha - \phi) - (1 - \alpha\phi) = -(1 - \alpha^2) < 0$.

Proof of Proposition 10. The proof has four steps.

Step 1. *Find the Wold representation of the observation vector $s_{it} = \Gamma(L)w_{it}$.* In any symmetric REE, $(1 - \rho L)\tilde{p}_t = A(L)v_t + \sigma_\eta(1 - \rho L)B(L)\eta_t$, for some one-sided $A(L)$ and $B(L)$.[16] The equilibrium law of motion of $s_{it} = (u_{it}, \tilde{p}_t)$ is then

$$s_{it} = \frac{1}{1-\rho L} \begin{bmatrix} \sigma_v & 0 & \sigma_\varepsilon(1-\rho L) \\ A(L) & \sigma_\eta(1-\rho L)B(L) & 0 \end{bmatrix} \begin{bmatrix} v_t \\ \eta_t \\ \varepsilon_{it} \end{bmatrix} \equiv \frac{1}{1-\rho L} M(L) e_{it}.$$

Given $M(L)$, the Wold factor $\Gamma(L)$ can be computed using the procedure from pp.44-

[16] $A(L)$ and $B(L)$ are re-scaled here relative to (18) only for analytical convenience.

47 of Rozanov (1967). The result is

$$\Gamma(L) = \frac{(1+\vartheta^2)^{-1/2}}{(1-\rho L)} \begin{bmatrix} \vartheta\sqrt{\frac{\rho}{\alpha}}\sigma_\varepsilon(L-\alpha) & -\sqrt{\frac{\rho}{\alpha}}\sigma_\varepsilon(1-\alpha L) \\ \vartheta\sqrt{\frac{\alpha}{\rho}\frac{\sigma_v}{\sigma_\varepsilon}\frac{A(L)L}{(1-\alpha L)}} + \gamma(L) & -\sqrt{\frac{\alpha}{\rho}\frac{\sigma_v}{\sigma_\varepsilon}\frac{A(L)L}{L-\alpha}} + \vartheta\frac{\gamma(L)(1-\alpha L)}{(L-\alpha)} \end{bmatrix}, \quad (40)$$

where

$$\vartheta \equiv \sqrt{\frac{\alpha}{\rho}}\frac{\sigma_v}{\sigma_\varepsilon}\frac{\alpha A(\alpha)}{(1-\alpha^2)\gamma(\alpha)}, \quad (41)$$

and $\gamma(L)$ is the univariate Wold factor that satisfies

$$\gamma(L)\gamma(L^{-1}) = \frac{\alpha}{\rho}\frac{(1-\rho L)(1-\rho L^{-1})}{(1-\alpha L)(1-\alpha L^{-1})}A(L)A(L^{-1}) + \sigma_\eta^2(1-\rho L)(1-\rho L^{-1})B(L)B(L^{-1}), \quad (42)$$

and α has the same definition as in the proof of Proposition 3.

Step 2. *Find the equilibrium fixed point* (19). According to the structural model,

$$(1-\rho L)\tilde{p}_t = \sigma_v v_t - \frac{b_1 f_0^2}{f_2}\frac{(1-\rho L)L}{(1-\lambda L)}\sum_{j=1}^{\infty}(\beta\lambda)^j \bar{E}_t u_{i,t+j} + \sigma_\eta(1-\rho L)\eta_t. \quad (43)$$

Using the Hansen and Sargent (1981) formula and averaging across sectors,

$$\sum_{j=1}^{\infty}(\beta\lambda)^j \bar{E}_t u_{i,t+j} = \frac{\beta\lambda}{L-\beta\lambda}\begin{bmatrix} 1 & 0 \end{bmatrix}(\Gamma(L)-\Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} \theta_t \\ \tilde{p}_t \end{bmatrix}.$$

Substituting this into (43) and matching coefficients on v_t and η_t implies

$$\begin{bmatrix} A(L) & \sigma_\eta(1-\rho L)B(L) \end{bmatrix} = \begin{bmatrix} \sigma_v & 0 \end{bmatrix} + \begin{bmatrix} 0 & \sigma_\eta(1-\rho L) \end{bmatrix}$$

$$-\frac{b_1 f_0^2}{f_2}\frac{\beta\lambda L}{(1-\lambda L)(L-\beta\lambda)}\begin{bmatrix} 1 & 0 \end{bmatrix}(\Gamma(L)-\Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} \sigma_v & 0 \\ A(L) & \sigma_\eta(1-\rho L)B(L) \end{bmatrix}.$$

Substituting in the expression for $\Gamma(L)$ in (40) and rearranging,

$$A(L) = \sigma_v \frac{(1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 L\psi(L)}{(1-\lambda L)(1-\alpha L)(L-\alpha) + b_0\vartheta\sigma_\varepsilon(1-\alpha^2)\sqrt{\frac{\alpha}{\rho}\frac{L(1-\rho L)^2(L-\rho)}{(1-\alpha L)\gamma(L)}}} \quad (44)$$

$$B(L) = \frac{\gamma(L)(1-\lambda L)(1-\alpha L)}{\gamma(L)(1-\lambda L)(1-\alpha L) + b_0(1-\alpha^2)\vartheta\sqrt{\frac{\rho}{\alpha}}\sigma_\varepsilon(1-\rho L)L}. \quad (45)$$

where $\psi(L)$ and b_0 have the same definitions as in the proof of Proposition 3.

Step 3. *Assume $A(L)$ and $B(L)$ are rational and re-write* (44) *and* (45) *in terms of polynomials.* If $A(L)$ and $B(L)$ are rational in L, then it is possible to write

$$A(L) = \sigma_A \frac{p_A(L)}{q_A(L)} \quad \text{and} \quad B(L) = \sigma_B \frac{p_B(L)}{q_B(L)} \tag{46}$$

where $p_i(L)$ and $q_i(L)$ are polynomials with no common zeros, $q_i(L)$ has no inside zeros, and $p_i(0) = q_i(0) = 1$ for $i = A, B$. In terms of these polynomials, (42) implies

$$\gamma(L) = \sigma_m \sqrt{\frac{\alpha}{\rho}} \frac{(1 - \rho L)m(L)}{(1 - \alpha L)q_A(L)q_B(L)}, \tag{47}$$

where $m(L)$ is a polynomial with no inside zeros, which satisfies

$$\begin{aligned} \sigma_m^2 m(L)m(L^{-1}) = {}& \sigma_A^2 p_A(L)p_A(L^{-1})q_B(L)q_B(L^{-1}) \\ & + \sigma_\eta^2 \frac{\rho}{\alpha}(1 - \alpha L)(1 - \alpha L^{-1})\sigma_B^2 p_B(L)p_B(L^{-1})q_A(L)q_A(L^{-1}), \end{aligned} \tag{48}$$

$m(0) = 1$, and $\sigma_m > 0$. Substituting (47) into (44) and (45), and rearranging,

$$\sigma_A \frac{p_A(L)}{q_A(L)} = \sigma_v \frac{m(L)(1 - \lambda L)(1 - \alpha L)(L - \alpha) - b_0 m(L) L \psi(L)}{m(L)(1 - \lambda L)(1 - \alpha L)(L - \alpha) + b_0 \vartheta \frac{\sigma_\varepsilon (1 - \alpha^2)}{\sigma_m} q_A(L)q_B(L)L(1 - \rho L)(L - \rho)} \tag{49}$$

$$\sigma_B \frac{p_B(L)}{q_B(L)} = \frac{m(L)(1 - \lambda L)}{m(L)(1 - \lambda L) + b_0(1 - \alpha^2)\vartheta \frac{\rho}{\alpha} \frac{\sigma_\varepsilon}{\sigma_m} L q_A(L) q_B(L)}. \tag{50}$$

Also, by (46), (47), and the definition of ϑ in (41),

$$\vartheta = \frac{\sigma_v}{\sigma_\varepsilon} \frac{\alpha}{1 - \alpha\rho} \frac{\sigma_A}{\sigma_m} \frac{p_A(\alpha)q_B(\alpha)}{m(\alpha)}. \tag{51}$$

Step 4. *Show that the fixed point equation defined by* (48), (49), (50), *and* (51) *has no solution, which proves that no rational $A(L)$ and $B(L)$ can solve (44) - (45).* The strategy is to derive a contradiction in each of six possible cases. First, it helps to simplify the equations and establish some preliminary results.

First, $\sigma_A = \sigma_v$ and $\sigma_B = 1$. This follows from evaluating (49) and (50) at $L = 0$, and using $p_A(0) = q_A(0) = p_B(0) = q_B(0) = 1$. Second, $\vartheta \neq 0$. If not, (49) implies

$$\frac{p_A(L)}{q_A(L)} = \frac{(1 - \lambda L)(1 - \alpha L) - b_0 L(\rho - \alpha)}{(1 - \lambda L)(1 - \alpha L)},$$

and (51) implies $p_A(\alpha) = 0$, so α must be a zero of the numerator on the right. But this is not possible, because (33) implies $b\alpha(\rho-\alpha) < (1-\alpha\lambda)(1-\alpha^2)$.

Define $\tilde{m}(L) \equiv m(L)(1-\lambda L)$, which has no inside zeros. Comparing numerators and denominators in (50), any zeros of $p_B(L)$ or $q_B(L)$ must be a zero of $\tilde{m}(L)$, so

$$\tilde{m}(L) = p_B(L)q_B(L)\tilde{m}_1(L) \tag{52}$$

for some polynomial $\tilde{m}_1(L)$. This implies $p_B(L)$ has no inside zeros. By (53), any zeros of $\tilde{m}_1(L)$ must be shared by $q_A(L)$; i.e.

$$q_A(L) = \tilde{m}_1(L)q_{A1}(L) \tag{53}$$

for some polynomial $q_{A1}(L)$. Using these definitions, (50) can be rewritten as

$$p_B(L) = q_B(L) - b_0\vartheta\frac{\sigma_\varepsilon}{\sigma_m}(1-\alpha^2)\frac{\rho}{\alpha}q_{A1}(L)L, \tag{54}$$

which implies that $q_B(L)$ and $q_{A1}(L)$ have no common zeros, since $q_B(L)$ and $p_B(L)$ do not. Substituting (52), (53), and (54) into (49), and rearranging,

$$\frac{p_A(L)(1-\lambda L)}{p_B(L)\tilde{m}_1(L)} = q_{A1}(L)\frac{N(L)}{D(L)}, \tag{55}$$

$$N(L) \equiv (1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 L[\vartheta^2(1-\alpha L)(1-\alpha\rho) + (L-\alpha)(\rho-\alpha)] \tag{56}$$

$$D(L) \equiv q_B(L)(1-\alpha L)(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)q_{A1}(L)L^2. \tag{57}$$

By (55), any inside zero of $D(L)$ must be shared by $N(L)$, because $\tilde{m}(L)$ and $q_{A1}(L)$ have no inside zeros, by definition, and $p_B(L)$ has no inside zeros, by (52). By (56), $\deg N(L) = 3$. Substituting (52), (53), and (55) into (48), and rearranging,

$$\begin{aligned} q_B(L)q_B(L^{-1})D(L)D(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2}q_{A1}(L)q_{A1}(L^{-1})q_B(L)q_B(L^{-1})N(L)N(L^{-1}) \\ &+\frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)(1-\alpha L^{-1})(1-\lambda L)(1-\lambda L^{-1})q_{A1}(L)q_{A1}(L^{-1})D(L)D(L^{-1}). \end{aligned} \tag{58}$$

This expression implies that if $q_{A1}(1/r) = 0$ for $|r| < 1$, then $(1-\alpha/r)D(r) = 0$. To see this, if $q_{A1}(1/r) = 0$ then $q_B(1/r) \neq 0$ because (54) implies $q_B(L)$ and $q_{A1}(L)$ have no common zeros, and $q_B(r) \neq 0$ because $q_B(L)$ has no inside zeros, so (58) implies $D(r)D(1/r) = 0$. But by (57), $D(1/r) = 0$ only if $r = \alpha$, which implies $(1-\alpha/r)D(r) = 0$.

Using theses preliminary results, it is possible to consider the six cases in the table below. The numbers in the table correspond to the numbers in the proof below.

$q_{A1}(1/\alpha) \neq 0$				$q_{A1}(1/\alpha) = 0$	
$q_B(1/\alpha)q_B(1/\lambda)\neq 0$	$q_B(1/\alpha)=0$ $q_B(1/\lambda)\neq 0$	$q_B(1/\alpha)\neq 0$ $q_B(1/\lambda)=0$	$q_B(1/\alpha)=0$ $q_B(1/\lambda)=0$	$q_B(1/\lambda)\neq 0$	$q_B(1/\lambda)=0$
1.1	1.2	1.3	1.4	2.1	2.2

Case 1. If $q_{A1}(1/\alpha) \neq 0$, then $q_{A1}(1/r) = 0$ implies $D(r) = 0$ by (58). This means $D(L)$ has an inside zero, which by (55) is shared by $N(L)$. But then a factor of the form $(L - r)(L^{-1} - r)$ cancels on both sides of (58), and the same argument can be repeated to show that the multiplicity of the zero r in $D(L)$ and $N(L)$ is arbitrarily large, contradicting $\deg N(L) = 3$. Therefore, $q_{A1}(L) = 1$, and (58) becomes

$$
\begin{aligned}
q_B(L)q_B(L^{-1})D(L)D(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2} q_B(L)q_B(L^{-1})N(L)N(L^{-1}) \\
&+ \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)(1-\alpha L^{-1})(1-\lambda L)(1-\lambda L^{-1})D(L)D(L^{-1}).
\end{aligned} \tag{59}
$$

By (59), if $q_B(1/r) = 0$, then $(1-\alpha/r)(1-\lambda/r)D(r) = 0$. Now consider each sub-case.

Case 1.1. If $q_B(1/\alpha)q_B(1/\lambda) \neq 0$, then $q_B(1/r) = 0$ implies $D(r) = 0$ since $q_{A1}(1/\alpha) \neq 0$ in this case, so $D(L)$ has an inside zero. But then the argument above can be used to produce a contradiction with $\deg N(L) = 3$. Therefore, $q_B(L) = 1$, and (57) and (59) become

$$
\begin{aligned}
D(L) &= (1-\alpha L)(L-\alpha) - b_0 \vartheta \frac{\sigma_v^2}{\sigma_\varepsilon \sigma_m}(1-\alpha^2)L^2 \\
D(L)D(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2} N(L)N(L^{-1}) \\
&\quad + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)(1-\alpha L^{-1})(1-\lambda L)(1-\lambda L^{-1})D(L)D(L^{-1}).
\end{aligned}
$$

If $\deg D(L) = 2$, then comparing degrees in the second equation implies the contradiction $2 = \max(3, 4)$. Therefore, the leading coefficient of $D(L)$ must vanish; i.e. $\alpha = -b_0\vartheta \frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)$, which implies $D(L) = (1+\alpha^2)L - \alpha$. Since $r \equiv \alpha/(1+\alpha^2) < \alpha$ is the only inside zero of $D(L)$, (55) implies $N(r) = 0$. By (56),

$$
\frac{\vartheta^2}{1+\vartheta^2} = -(\alpha - r)\frac{(1-\lambda r)(1-\alpha r) - br(\rho - \alpha)}{br[(1-\alpha r)(1-\alpha\rho) + (\alpha - r)(\rho - \alpha)]}.
$$

But by (33) and $r < \alpha$, it follows that $b < 1 - \lambda r$, so $(1-\lambda r)(1-\alpha r) - br(\rho - \alpha) > (1-\lambda r)(1-\rho r) > 0$. Therefore $\vartheta^2 < 0$, which is a contradiction.

Case 1.2. If $q_B(1/\alpha) = 0$ but $q_B(1/\lambda) \neq 0$, then $q_B(L) = (1-\alpha L)q_{B1}(L)$ for some polynomial $q_{B1}(L)$. Then (59) becomes

$$\begin{aligned} q_{B1}(L)q_{B1}(L^{-1})D(L)D(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2} q_{B1}(L)q_{B1}(L^{-1})N(L)N(L^{-1}) \\ &+ \frac{\sigma_\eta^2}{\sigma_m^2}(1-\lambda L)(1-\lambda L^{-1})D(L)D(L^{-1}). \end{aligned} \tag{60}$$

By (60), $q_{B1}(1/r) = 0$ implies $D(r) = 0$, and $D(L)$ has an inside zero, which can be used to produce a contradiction with $\deg N(L) = 3$. Therefore, $q_{B1}(L) = 1$, and (57) and (59) become

$$D(L) = (1-\alpha L)^2(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)L^2$$
$$D(L)D(L^{-1}) = \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1}) + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\lambda L)(1-\lambda L^{-1})D(L)D(L^{-1}).$$

From the first equation, $\deg D(L) = 3$. Comparing degrees on both sides of the second equation produces the contradiction $3 = \max(3,4)$.

Case 1.3. If $q_B(1/\lambda) = 0$ but $q_B(1/\alpha) \neq 0$, it is possible to write $q_B(L) = (1-\lambda L)q_{B1}(L)$ for some polynomial $q_{B1}(L)$. Then (59) becomes

$$\begin{aligned} q_{B1}(L)q_{B1}(L^{-1})D(L)D(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2} q_{B1}(L)q_{B1}(L^{-1})N(L)N(L^{-1}) \\ &+ \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)(1-\alpha L^{-1})D(L)D(L^{-1}). \end{aligned} \tag{61}$$

By (61), $q_{B1}(1/r) = 0$ implies $D(r) = 0$, and $D(L)$ has an inside zero, which can be used in (61) to produce a contradiction with $\deg N(L) = 3$. Therefore, $q_{B1}(L) = 1$, and (57) and (61) become

$$D(L) = (1-\lambda L)(1-\alpha L)(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)L^2$$
$$D(L)D(L^{-1}) = \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1}) + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)(1-\alpha L^{-1})D(L)D(L^{-1}).$$

From the first equation, $\deg D(L) = 3$. Comparing degrees on both sides of the second equation produces the contradiction $3 = \max(3,4)$.

Case 1.4. If $q_B(1/\lambda) = 0$ and $q_B(1/\alpha) = 0$, it is possible to write $q_B(L) = (1-\alpha L)(1-\lambda L)q_{B1}(L)$ for some polynomial $q_{B1}(L)$. Then (59) becomes

$$q_{B1}(L)q_{B1}(L^{-1})D(L)D(L^{-1}) = \frac{\sigma_v^2}{\sigma_m^2}q_{B1}(L)q_{B1}(L^{-1})N(L)N(L^{-1}) + \frac{\sigma_\eta^2}{\sigma_m^2}D(L)D(L^{-1}).$$

This implies $q_{B1}(L) = 1$; otherwise it can be shown that the reciprocal of any zero of $q_{B1}(L)$ would be an inside zero of $D(L)$ and $N(L)$ of multiplicity greater than 3. So,

$$D(L) = (1-\lambda L)(1-\alpha L)^2(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)L^2$$
$$D(L)D(L^{-1}) = \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1}) + \frac{\sigma_\eta^2}{\sigma_m^2}D(L)D(L^{-1}).$$

The first equation implies $\deg D(L) = 4$, and the second equation implies all the zeros of $D(L)$ must cancel with zeros of $N(L)$. But $\deg N(L) = 3$, so this is a contradiction.

Case 2. If $q_{A1}(1/\alpha) = 0$, it is possible to write $q_{A1}(L) = (1-\alpha L)q_{A2}(L)$ for some polynomial $q_{A2}(L)$. Then $D(L) = (1-\alpha L)D_1(L)$, where

$$D_1(L) = q_B(L)(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)q_{A2}(L)L^2. \tag{62}$$

By (58), $q_{A2}(1/r) = 0$ implies $D_1(r) = 0$, which can be used to produce a contradiction with $\deg N(L) = 3$. Therefore, $q_{A2}(L) = 1$ and (58) becomes

$$\begin{aligned} q_B(L)q_B(L^{-1})D_1(L)D_1(L^{-1}) = {} & \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1})q_B(L)q_B(L^{-1}) \\ & + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)^2(1-\alpha L^{-1})^2(1-\lambda L)(1-\lambda L^{-1})D_1(L)D_1(L^{-1}). \end{aligned} \tag{63}$$

By (63), $q_B(1/r) = 0$ implies $(1-\lambda/r)D_1(r) = 0$. Now consider each sub-case.

Case 2.1. If $q_B(1/\lambda) \neq 0$, then $q_B(1/r) = 0$ implies $D_1(r) = 0$, so $D_1(L)$ has an inside zero, which can be used to produce a contradiction with $\deg N(L) = 3$.

Therefore, $q_B(L) = 1$, and (62) and (63) become

$$\begin{aligned} D_1(L) &= (L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)L^2 \\ D_1(L)D_1(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1}) \\ &\quad + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)^2(1-\alpha L^{-1})^2(1-\lambda L)(1-\lambda L^{-1})D_1(L)D_1(L^{-1}). \end{aligned} \tag{64}$$

From the first equation, $\deg D_1(L) = 2$. Comparing degrees on both sides of the second equation produces the contradiction $2 = \max(3,5)$.

Case 2.2. If $q_B(1/\lambda) = 0$, then $q_B(L) = (1-\lambda L)q_{B1}(L)$ for some polynomial $q_{B1}(L)$, and (63) becomes

$$\begin{aligned} q_{B1}(L)q_{B1}(L^{-1})D_1(L)D_1(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2}q_{B1}(L)q_{B1}(L^{-1})N(L)N(L^{-1}) \\ &\quad + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)^2(1-\alpha L^{-1})^2D_1(L)D_1(L^{-1}). \end{aligned} \tag{65}$$

If $q_{B1}(L) \neq 1$, then this equation implies $D_1(L)$ has an inside zero, which can be used to show that $N(L)$ has an inside zero of arbitrarily large multiplicity, which is not possible. Therefore, it must be the case that $q_{B1}(L) = 1$, so (57) and (65) become

$$\begin{aligned} D_1(L) &= (1-\lambda L)(L-\alpha) - b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)L^2 \\ D_1(L)D_1(L^{-1}) &= \frac{\sigma_v^2}{\sigma_m^2}N(L)N(L^{-1}) + \frac{\sigma_\eta^2}{\sigma_m^2}(1-\alpha L)^2(1-\alpha L^{-1})^2D_1(L)D_1(L^{-1}). \end{aligned}$$

If $\deg D_1(L) = 2$, then the second equation implies the contradiction $2 = \max(3,4)$. Therefore, the leading coefficient of $D_1(L)$ must vanish, i.e. $\lambda = -b_0\vartheta\frac{\sigma_v^2}{\sigma_\varepsilon\sigma_m}(1-\alpha^2)$, which implies $D_1(L) = (1+\alpha\lambda)L - \alpha$. Therefore, $\deg D_1(L) = 1$, and its zero is $r \equiv \alpha/(1+\alpha\lambda) < \alpha$. Since this is an inside zero, (55) implies $N(r) = 0$. Using (56),

$$\frac{\vartheta^2}{1+\vartheta^2} = -(\alpha - r)\frac{(1-\lambda r)(1-\alpha r) - br(\rho-\alpha)}{br[(1-\alpha r)(1-\alpha\rho) + (\alpha - r)(\rho-\alpha)]}.$$

But by (33) and $r < \alpha$, $(1-\lambda r)(1-\alpha r) - br(\rho-\alpha) > (1-\lambda r)(1-\rho r) > 0$. Therefore $\vartheta^2 < 0$, which is a contradiction.

Online Appendix to:

Revisiting the Forecasts of Others

Ryan Chahrour — Boston College
Kyle Jurado — Duke University

Proof of Proposition 4. The fact that the FCE paths are a REE follows from the fact that the operator on the right side of (7) is invertible into the past, as shown in the proof of Proposition 2. What remains is to show that this REE is unique. In any REE, $k_{i,t+1}$ is measurable with respect to $s_i^t = \bar{p}^t$, so $k_{i,t+1} = A_i(L)\bar{p}_t$ for some one-sided operator $A_i(L)$. Substituting this expression for $k_{i,t+1}$ into (2) and averaging across i implies

$$\left[1 + \frac{1}{n}\sum_{i=1}^{n} b_1 f_0 L A_i(L)\right]\bar{p}_t = \bar{u}_t,$$

so $\bar{u}_t$ is measurable with respect to $\bar{p}^t$. And since $\bar{p}_t$ cannot be correlated with future disturbances, the operator on the left must be invertible, implying that $\bar{p}_t$ is also measurable with respect to $\bar{u}^t$. Therefore, $\bar{p}^t$ contains the same information as $\bar{u}^t$ in any REE.

Proof of Proposition 5. The existence and uniqueness of the FCE follows from the same reasoning as in the proof of Proposition 1, and the policy function (6) is the same. Given that policy function, the closed-form expression in the proposition comes from evaluating the forecasts $E_{it}u_{i,t+j}$ under the new law of motion (10). To do so, first notice that (10) implies

$$E_{it}u_{i,t+j} = \rho_\theta^j E_{it}\theta_t + \rho_z^j E_{it}z_{it} = (\rho_\theta^j - \rho_z^j)E_{it}\theta_t + \rho_z^j u_{it}, \tag{66}$$

where the second equality uses $z_{it} = u_{it} - \theta_t$ to substitute out z_{it}. The vector $u_t = (u_{1t}, \ldots, u_{nt})$ contains the same information as $(\bar{u}_t, \hat{u}_{1t}, \ldots, \hat{u}_{2t})$, where $\hat{u}_{it} \equiv u_{it} - \bar{u}_t$, since the two are related by a non-degenerate linear transformation. Since the processes $\{\hat{u}_{1t}, \ldots, \hat{u}_{nt}\}$ are independent of $\{\theta_t\}$,

$$E_{it}\theta_t = E(\theta_t|u^t) = E(\theta_t|\bar{u}^t, \hat{u}_1^t, \ldots, \hat{u}_n^t) = E(\theta_t|\bar{u}^t). \tag{67}$$

Now it is necessary to compute $E(\theta_t|\bar{u}^t)$. By (10),

$$\bar{u}_t = \theta_t + \sigma_\varepsilon \bar{\varepsilon}_t, \qquad \theta_t = \rho_\theta \theta_{t-1} + \sigma_v v_t, \qquad \bar{z}_t = \rho_z \bar{z}_{t-1} + \sigma_\varepsilon \bar{\varepsilon}_t,$$

where $\bar{\varepsilon}_t \equiv \frac{1}{n}\sum_{i=1}^n \varepsilon_t$ is white noise with variance $1/n$. This implies that the Wold factor associated with the spectral density of $\{\bar{u}_t\}$ is

$$H(L) = \sqrt{\frac{\rho_\theta \sigma_\varepsilon^2 + \rho_z n \sigma_v^2}{\phi n}} \frac{(1-\phi L)}{(1-\rho_\theta L)(1-\rho_z L)},$$

where ϕ is defined in the proposition. By the Wiener-Kolmogorov filtering equation,

$$E(\theta_t|\bar{u}^t) = \left[\frac{\sigma_v^2}{(1-\rho_\theta L)(1-\rho_\theta L^{-1})} H(L^{-1})^{-1}\right]_+ H(L)^{-1}\bar{u}_t = \frac{(\rho_\theta - \phi)}{(\rho_\theta - \rho_z)} \frac{(1-\rho_z L)}{(1-\phi L)} \bar{u}_t, \tag{68}$$

where $[\ \cdot\ ]_+$ projects onto the space spanned by non-negative powers of L, and the second equality uses the fact that, by definition of ϕ,

$$\frac{\phi n \sigma_v^2}{\rho_\theta \sigma_\varepsilon^2 + \rho_z n \sigma_v^2} \frac{(1-\rho_\theta \rho_z)}{(1-\rho_\theta \phi)} = \frac{(\rho_\theta - \phi)}{(\rho_\theta - \rho_z)}.$$

Substituting (66), (67), and (68) into (6) delivers the final expression for $k_{i,t+1}$.

Proof of Proposition 6. By substituting the closed-form expression (11) into the demand curve (2) and averaging across sectors,

$$\bar{p}_t = \left[1 - \frac{b_1 f_0 L}{(1-\lambda L)(1-\phi L)} \Big(\omega_\theta (1-\rho_z L) + \omega_z (1-\phi L)\Big)\right] \bar{u}_t. \tag{69}$$

To prove the proposition, it is sufficient to prove that the operator on the right side of this equation is invertible into the past. This holds if and only if

$$\mathcal{P}(\mu) = (\mu - \lambda)(\mu - \phi) - b_1 f_0 (\omega_\theta(\mu - \rho_z) + \omega_z(\mu - \phi))$$

has no outside zeros.

First, suppose that $\rho_\theta = \rho_z$. Then the characteristic equation simplifies to

$$\mathcal{P}(\mu) = (\mu - \phi)\left(\mu - \lambda - \frac{b_1 f_0^2 \beta\lambda\rho_z}{f_2(1-\beta\lambda\rho_z)}\right),$$

which has two inside zeros, because

$$0 < \lambda + \frac{b_1 f_0^2 \beta\lambda\rho_z}{f_2(1-\beta\lambda\rho_z)} < \lambda + \frac{(b_1 f_0^2 + f_1)\beta\lambda}{f_2(1-\beta\lambda)} = 1$$

by definition of λ.

Second, suppose that $\rho_\theta \neq \rho_z$. Then

$$\mathcal{P}(0) = \lambda\phi + \frac{b_1 f_0^2}{f_2}\frac{\beta\lambda\rho_\theta\rho_z(1-\beta\lambda\phi)}{(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)} > 0,$$
$$\mathcal{P}(\phi) = -\frac{b_1 f_0^2}{f_2}\frac{\beta\lambda(\rho_\theta-\phi)(\phi-\rho_z)}{(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)} < 0,$$

and

$$\mathcal{P}(1) = (1-\lambda)(1-\phi) - \frac{b_1 f_0^2}{f_2}\frac{\beta\lambda}{(1-\beta\lambda\rho_z)}\left[\frac{(\rho_\theta-\phi)(1-\rho_z)}{(1-\beta\lambda\rho_\theta)} + \rho_z(1-\phi)\right]$$
$$> (1-\phi)\left[(1-\lambda) - \frac{b_1 f_0^2 + f_1}{f_2}\frac{\beta\lambda(1-\beta\lambda\rho_\theta\rho_z)}{(1-\beta\lambda\rho_z)(1-\beta\lambda\rho_\theta)}\right] \qquad (f_1 \geq 0,\ \rho_\theta < 1)$$
$$= (1-\phi)\frac{b_1 f_0^2 + f_1}{f_2}\frac{\beta^2\lambda^2(1-\rho_\theta)(1-\rho_z)}{(1-\beta\lambda)(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)} > 0. \quad \left(1-\lambda = \tfrac{b_1 f_0^2 + f_1}{f_2}\tfrac{\beta\lambda}{1-\beta\lambda}\right)$$

Therefore, $\mathcal{P}(\mu)$ has two inside zeros in this case as well.

Proof of Proposition 7. The proof has four steps.

Step 1. *Prove that $\bar{p}_t = A(L)\bar{u}_t$ in any symmetric REE, with $A(L)$ one-sided into the past.* The proof of this step is the same as in the proof of Proposition 3.

Step 2. *Find the Wold representation of the observation vector $s_{it} = \Gamma(L)w_{it}$.* The law of motion for s_{it} is

$$s_{it} = \frac{1}{(1-\rho_\theta L)(1-\rho_z L)}\begin{bmatrix} \sigma_v(1-\rho_z L) & \sigma_\varepsilon(1-\rho_\theta L)\iota_i' \\ \sigma_v(1-\rho_z L)A(L) & \sigma_\varepsilon(1-\rho_\theta L)A(L)\frac{1}{n}1_n' \end{bmatrix}\begin{bmatrix} v_t \\ \varepsilon_t \end{bmatrix}$$
$$\equiv \frac{1}{(1-\rho_\theta L)(1-\rho_z L)}M_i(L)e_t,$$

where ι_i is a vector of zeros with a one in the i-th position, and 1_n is an n-dimensional vector of ones. Given $M_i(L)$, the Wold factor $\Gamma(L)$ can be computed using the procedure from pp.44-47 of Rozanov (1967). The result is

$$\Gamma(L) = \frac{(1+\vartheta^2)^{-1/2}}{(1-\rho_\theta L)(1-\rho_z L)} \times$$

$$\begin{bmatrix} \vartheta\sqrt{\frac{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}{\alpha}}\sigma_\varepsilon(L-\alpha) & -\sqrt{\frac{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}{\alpha}}\sigma_\varepsilon(1-\alpha L) \\ \vartheta\sqrt{\frac{\alpha}{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}}\frac{\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2}{\phi n}\frac{(1-\phi L)(L-\phi)}{(1-\alpha L)}A(L)+\gamma(L) & -\sqrt{\frac{\alpha}{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}}\frac{\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2}{\phi n}\frac{(1-\phi L)(L-\phi)}{(L-\alpha)}A(L)+\vartheta\gamma(L)\frac{1-\alpha L}{L-\alpha} \end{bmatrix}, \tag{70}$$

where

$$\vartheta \equiv \sqrt{\frac{\alpha}{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}}\frac{\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2}{\phi n}\frac{(1-\alpha\phi)(\alpha-\phi)}{(1-\alpha^2)}\frac{A(\alpha)}{\gamma(\alpha)}, \tag{71}$$

$\gamma(L)$ is the univariate Wold factor that satisfies

$$\gamma(L)\gamma(L^{-1}) = \sigma_\varepsilon^2(n-1)\frac{\alpha(\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2)}{\phi(\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2)}\frac{(1-\phi L)(1-\phi L^{-1})(1-\rho_\theta L)(1-\rho_\theta L^{-1})}{(1-\alpha L)(1-\alpha L^{-1})} \times A(L)A(L^{-1}), \tag{72}$$

α is the smaller zero of

$$(\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2)\alpha^2 - \left(\sigma_\varepsilon^2(1+\rho_\theta^2)+\sigma_v^2(1+\rho_z^2)\right)\alpha + (\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2) = 0, \tag{73}$$

and ϕ is the smaller zero of

$$(\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2)\phi^2 - \left(\sigma_\varepsilon^2(1+\rho_\theta^2)+n\sigma_v^2(1+\rho_z^2)\right)\phi + (\rho_\theta\sigma_\varepsilon^2+\rho_z n\sigma_v^2) = 0. \tag{74}$$

By definition, the parameters α and ϕ satisfy the relations

$$\rho_z < \phi < \alpha < \rho_\theta \quad \text{if} \quad \rho_\theta > \rho_z \quad \text{and} \quad \rho_z > \phi > \alpha > \rho_\theta \quad \text{if} \quad \rho_\theta < \rho_z. \tag{75}$$

Step 3. *Find the equilibrium fixed point equation* $A(L) = T[A(L)]$. Using the

Hansen and Sargent (1981) formula and averaging across sectors,

$$\sum_{j=1}^{\infty}(\beta\lambda)^j \bar{E}_t u_{i,t+j} = \frac{\beta\lambda}{L-\beta\lambda}\left[\,1 \quad 0\,\right](\Gamma(L)-\Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} 1 \\ A(L) \end{bmatrix}\bar{u}_t.$$

Substituting this into (9) and matching coefficients on $\bar{u}_t$ implies

$$A(L) = 1 - \frac{b_1 f_0^2}{f_2}\frac{\beta\lambda L}{(1-\lambda L)(L-\beta\lambda)}\left[\,1 \quad 0\,\right](\Gamma(L)-\Gamma(\beta\lambda))\Gamma(L)^{-1}\begin{bmatrix} 1 \\ A(L) \end{bmatrix}.$$

Substituting in the expression for $\Gamma(L)$ in (70) and rearranging,

$$A(L) = \frac{(1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 L\psi(L)}{(1-\lambda L)(1-\alpha L)(L-\alpha) + b_0\vartheta\frac{\sigma_\varepsilon^2(n-1)(1-\alpha^2)}{n}\sqrt{\frac{\alpha}{\rho_\theta\sigma_\varepsilon^2+\rho_z\sigma_v^2}}\frac{(1-\rho_z L)(1-\rho_\theta L)^2(L-\rho_\theta)L}{\gamma(L)(1-\alpha L)}}, \tag{76}$$

where

$$\psi(L) \equiv \vartheta^2(1-\alpha L)\psi_1(L) + (L-\alpha)\psi_2(L), \tag{77}$$

$$\psi_1(L) \equiv 1 - \alpha(\rho_\theta + \rho_z - \rho_\theta\rho_z\beta\lambda) + \rho_\theta\rho_z(\alpha-\beta\lambda)L, \tag{78}$$

$$\psi_2(L) \equiv \rho_\theta + \rho_z - \alpha - \rho_\theta\rho_z\beta\lambda - \rho_\theta\rho_z(1-\alpha\beta\lambda)L, \tag{79}$$

$b_0 \equiv b/(1+\vartheta^2)$, and

$$b \equiv \frac{b_1 f_0^2\beta\lambda}{f_2(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)} > 0.$$

An important property of the parameter b is that

$$b < \frac{(1-\lambda)(1-\beta\lambda)}{(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z)}. \tag{80}$$

To see this, note that by the definition of λ,

$$(1-\lambda)(1-\beta\lambda) = \frac{(b_1 f_0^2 + f_1)\beta\lambda}{f_2} > \frac{b_1 f_0^2\beta\lambda}{f_2} = b(1-\beta\lambda\rho_\theta)(1-\beta\lambda\rho_z).$$

Step 4. *Prove that $A(L)$ must be invertible.* First consider the case when $\rho_\theta = \rho_z \equiv \rho$. Equations (73) and (74) imply that $\alpha = \phi = \rho$. By (71), this implies that

$\vartheta = 0$, so (76) becomes

$$A(L) = \frac{(1-\lambda L)(1-\alpha L) - bL[\rho - \rho^2(\beta\lambda + (1-\alpha\beta\lambda)L)]}{(1-\lambda L)(1-\alpha L)}$$

$$= \frac{1 - [\lambda + b\rho(1-\rho\beta\lambda)]L}{1-\lambda L}.$$

This operator is invertible because the moving average coefficient in the numerator is no greater than one:

$$\lambda + b\rho(1-\rho\beta\lambda) \leq \lambda + \frac{\rho(1-\lambda)(1-\beta\lambda)}{(1-\rho\beta\lambda)} = \frac{\lambda(1-\rho)+\rho-\rho\beta\lambda}{1-\rho\beta\lambda} \leq 1,$$

where the first inequality uses (80) and the second uses $0 < \lambda < 1$. Therefore, for the remainder of the proof it can be assumed that $\rho_\theta \neq \rho_z$.

Second, note that when $\rho_\theta \neq \rho_z$ it must be that $\vartheta \neq 0$. To see this, suppose to the contrary that $\vartheta = 0$, so that (76) becomes

$$A(L) = \frac{(1-\lambda L)(1-\alpha L) - bL\psi_2(L)}{(1-\lambda L)(1-\alpha L)}.$$

By (71), $\vartheta = 0$ implies $A(\alpha) = 0$, since $\alpha \neq \phi$ and $\gamma(L)$ has no inside zeros. Therefore, the polynomial in the numerator on the right side of this equation must have a zero at α. But this is not possible, because (79) and (80) imply that

$$b\alpha\psi_2(\alpha) < (1-\alpha\lambda)(1-\alpha^2).$$

Third, let $H(L)$ denote the univariate Wold factor that satisfies

$$H(L)H(L^{-1}) = A(L)A(L^{-1}). \tag{81}$$

By (30),

$$\gamma(L) = \sqrt{\frac{\sigma_\varepsilon^2}{n}\frac{\alpha}{\phi}\left(1-\frac{1}{n}\right)\frac{(\rho_\theta\sigma_\varepsilon^2 + \rho_z\sigma_v^2)}{(\rho_\theta\sigma_\varepsilon^2 + \rho_z n\sigma_v^2)}}\,\frac{(1-\phi L)(1-\rho_\theta L)}{(1-\alpha L)}H(L). \tag{82}$$

Substitution of this expression into (76) produces

$$A(L) = \frac{H(L)(1-\lambda L)(1-\alpha L)(L-\alpha) - b_0 H(L)L\psi(L)}{H(L)(1-\lambda L)(1-\alpha L)(L-\alpha) + b_0(1-\alpha^2)\vartheta\sigma\frac{(1-\rho_z L)(1-\rho_\theta L)(L-\rho_\theta)L}{(1-\phi L)}}, \tag{83}$$

where $\sigma^2 \equiv (n-1)\phi\sigma_\varepsilon^2/(\rho_\theta\sigma_\varepsilon^2 + \rho_z n\sigma_v^2)$. By definition of α and ϕ,

$$\frac{\rho_\theta\sigma_\varepsilon^2 + \rho_z\sigma_v^2}{\alpha}(1-\alpha L)(L-\alpha) = \sigma_v^2(1-\rho_z L)(L-\rho_z) + \sigma_\varepsilon^2(1-\rho_\theta L)(L-\rho_\theta),$$

$$\frac{\rho_\theta\frac{\sigma_\varepsilon^2}{n} + \rho_z\sigma_v^2}{\phi}(1-\phi L)(L-\phi) = \sigma_v^2(1-\rho_z L)(L-\rho_z) + \frac{\sigma_\varepsilon^2}{n}(1-\rho_\theta L)(L-\rho_\theta).$$

Evaluating these equations at $L = \alpha$ and subtracting the first from the second, σ^2 can be written more conveniently as

$$\sigma^2 = \frac{(1-\alpha\phi)(\alpha-\phi)}{(1-\alpha\rho_\theta)(\rho_\theta-\alpha)}. \tag{84}$$

Now, suppose to the contrary that $A(L)$ is not invertible, and has at least one inside zero (which is not at the origin, since $A(0) = 1$). By (81), this means that $H(L)$ has at least one outside zero that is not shared by $A(L)$. By (83), there are only three possibilities:

1. $H(L) = \frac{1-\rho_\theta L}{L-\rho_\theta}A(L)$
2. $H(L) = \frac{1-\rho_z L}{L-\rho_z}A(L)$
3. $H(L) = \frac{(1-\rho_\theta L)(1-\rho_z L)}{(L-\rho_\theta)(L-\rho_z)}A(L)$.

These are the only possibilities because if $h(1/r) = 0$ and $A(1/r) \neq 0$ for any other $|r| < 1$, then $1/r$ would be a zero of the numerator but not the denominator of (83). What remains is to show that each of these three possibilities entails a contradiction.

Case 1. $H(L) = \frac{1-\rho_\theta L}{L-\rho_\theta}A(L)$. Substituting this expression for $H(L)$ into the fixed point equation (83) and solving for $A(L)$ implies that

$$A(L) = 1 - \frac{b_0 L[(1-\phi L)\psi(L) + (1-\alpha^2)\vartheta\sigma(1-\rho_z L)(L-\rho_\theta)^2]}{(1-\lambda L)(1-\alpha L)(L-\alpha)(1-\phi L)}. \tag{85}$$

Equation (85) and the hypothesis that $A(\rho_\theta) = 0$ provide an expression for ϑ^2,

$$\vartheta^2 = -\frac{(\rho_\theta-\alpha)}{(1-\alpha\rho_\theta)}\frac{(1-\lambda\rho_\theta)(1-\alpha\rho_\theta) - b\rho_\theta\psi_2(\rho_\theta)}{(1-\lambda\rho_\theta)(\rho_\theta-\alpha) - b\rho_\theta\psi_1(\rho_\theta)}, \tag{86}$$

where $\psi_1(L)$ and $\psi_2(L)$ are defined in (78) and (79), and $b_0 = b/(1+\vartheta^2)$ has been used to substitute out b_0. If $\rho_\theta < \alpha$, then by (78) and (80), and (79) and (80), respectively,

$$b\rho_\theta\psi_1(\rho_\theta) > (1-\lambda\rho_\theta)(\rho_\theta-\alpha) \quad \text{and} \quad b\rho_\theta\psi_2(\rho_\theta) < (1-\lambda\rho_\theta)(1-\alpha\rho_\theta). \tag{87}$$

By (86) and (87), $\vartheta^2 > 0$ implies $\rho_\theta > \alpha$, which implies $\rho_z < \phi < \alpha < \rho_\theta$, by (75).

Since $A(L)$ must be one-sided, the numerator of the fraction on the right side of (85) must have a zero at $L = \alpha$ to cancel the factor $L - \alpha$ in the denominator. This provides a second expression for ϑ^2,

$$\vartheta^2 = \frac{(\alpha - \phi)(\rho_\theta - \alpha)^3}{(1 - \alpha\phi)(1 - \alpha\rho_\theta)^3}, \tag{88}$$

where (77) and (84) have been used. Equating the two expressions for ϑ^2 in (86) and (88), it follows that

$$\frac{(\alpha - \phi)(\rho_\theta - \alpha)^2}{(1 - \alpha\phi)(1 - \alpha\rho_\theta)^2} = -\frac{(1 - \lambda\rho_\theta)(1 - \alpha\rho_\theta) - b\rho_\theta\psi_2(\rho_\theta)}{(1 - \lambda\rho_\theta)(\rho_\theta - \alpha) - b\rho_\theta\psi_1(\rho_\theta)}.$$

Since (80) implies $b(1 - \beta\lambda\rho_z) < (1 - \lambda\rho_\theta)$, this equation implies

$$\begin{aligned}\frac{(\alpha - \phi)(\rho_\theta - \alpha)^2}{(1 - \alpha\phi)(1 - \alpha\rho_\theta)^2} &> -\frac{(1 - \beta\lambda\rho_z)(1 - \alpha\rho_\theta) - \rho_\theta\psi_2(\rho_\theta)}{(1 - \beta\lambda\rho_z)(\rho_\theta - \alpha) - \rho_\theta\psi_1(\rho_\theta)} \\ &= \frac{(1 - \rho_\theta\rho_z) - \beta\lambda\rho_z(1 - \alpha\rho_\theta)}{\alpha(1 - \rho_\theta\rho_z) + \beta\lambda\rho_z(\rho_\theta - \alpha)},\end{aligned}$$

where the equality uses (78) and (79). Since $(\rho_\theta - \alpha) < (1 - \alpha\rho_\theta)$ and $\beta\lambda < 1$, this inequality implies

$$\frac{(\alpha - \phi)}{(1 - \alpha\phi)} > \frac{(1 - \rho_\theta\rho_z) - \rho_z(1 - \alpha\rho_\theta)}{\alpha(1 - \rho_\theta\rho_z) + \rho_z(\rho_\theta - \alpha)}.$$

But this is a contradiction, because the denominators on both sides are positive, and

$$\begin{aligned}(\alpha - \phi)[\alpha(1 - \rho_\theta\rho_z) + \rho_z(\rho_\theta - \alpha)] - (1 - \alpha\phi)[(1 - \rho_\theta\rho_z) - \rho_z(1 - \alpha\rho_\theta)] \\ = (1 - \alpha^2)[-1 + \rho_z + \rho_\theta\rho_z(1 - \phi)] \\ < (1 - \alpha^2)(\rho_z - \phi) < 0.\end{aligned}$$

Case 2. $H(L) = \frac{1-\rho_z L}{L-\rho_z}A(L)$. Substituting this expression for $H(L)$ into the fixed point equation (83) and solving for $A(L)$ implies that

$$A(L) = 1 - \frac{b_0 L[(1 - \phi L)\psi(L) + (1 - \alpha^2)\vartheta\sigma(1 - \rho_\theta L)(L - \rho_\theta)(L - \rho_z)]}{(1 - \lambda L)(1 - \alpha L)(L - \alpha)(1 - \phi L)}, \tag{89}$$

where σ^2 is still defined as in (84). Equation (89) and the hypothesis that $A(\rho_z) = 0$

provide an expression for ϑ^2,

$$\vartheta^2 = -\frac{(\rho_z - \alpha)}{(1-\alpha\rho_z)} \frac{(1-\lambda\rho_z)(1-\alpha\rho_z) - b\rho_z\psi_2(\rho_z)}{(1-\lambda\rho_z)(\rho_z-\alpha) - b\rho_z\psi_1(\rho_z)}. \tag{90}$$

By (78), (79), and (80), if $\rho_z < \alpha$ then

$$b\rho_z\psi_2(\rho_z) < (1-\lambda\rho_z)(1-\alpha\rho_z) \quad \text{and} \quad b\rho_z\psi_1(\rho_z) > (1-\lambda\rho_z)(\rho_z-\alpha). \tag{91}$$

Therefore, (90) indicates that $\vartheta^2 > 0$ only if $\rho_z > \alpha$, which implies $\rho_z > \phi > \alpha > \rho_\theta$ by (75). Because $A(L)$ must be one-sided, the numerator of the fraction on the right side of (89) must have a zero at $L = \alpha$ to cancel the factor $L - \alpha$ in the denominator. This provides a second expression for ϑ^2,

$$\vartheta^2 = \frac{(\alpha-\phi)(\rho_\theta-\alpha)(\rho_z-\alpha)^2}{(1-\alpha\phi)(1-\alpha\rho_\theta)(1-\alpha\rho_z)^2}, \tag{92}$$

where (77) and (84) have been used. Equating the expressions for ϑ^2 in (90) and (92),

$$\frac{(\alpha-\phi)(\rho_\theta-\alpha)(\rho_z-\alpha)}{(1-\alpha\phi)(1-\alpha\rho_\theta)(1-\alpha\rho_z)} = -\frac{(1-\lambda\rho_z)(1-\alpha\rho_z) - b\rho_z\psi_2(\rho_z)}{(1-\lambda\rho_z)(\rho_z-\alpha) - b\rho_z\psi_1(\rho_z)}.$$

Since (80) implies $b(1-\beta\lambda\rho_\theta) < (1-\lambda\rho_z)$, this equation implies

$$\begin{aligned}\frac{(\alpha-\phi)(\rho_\theta-\alpha)(\rho_z-\alpha)}{(1-\alpha\phi)(1-\alpha\rho_\theta)(1-\alpha\rho_z)} &> -\frac{(1-\beta\lambda\rho_\theta)(1-\alpha\rho_z) - \rho_z\psi_2(\rho_z)}{(1-\beta\lambda\rho_\theta)(\rho_z-\alpha) - \rho_z\psi_1(\rho_z)} \\ &= \frac{(1-\rho_\theta\rho_z) - \beta\lambda\rho_\theta(1-\alpha\rho_z)}{\alpha(1-\rho_\theta\rho_z) + \beta\lambda\rho_\theta(\rho_z-\alpha)},\end{aligned}$$

where the equality uses (78) and (79). Since $(\rho_\theta - \alpha) < (1-\alpha\rho_\theta)$, $(\rho_z - \alpha) < (1-\alpha\rho_z)$, and $\beta\lambda < 1$, this inequality implies

$$\frac{(\alpha-\phi)}{(1-\alpha\phi)} > \frac{(1-\rho_\theta\rho_z) - \rho_\theta(1-\alpha\rho_z)}{\alpha(1-\rho_\theta\rho_z) + \rho_\theta(\rho_z-\alpha)}.$$

But this is a contradiction, because the denominators on both sides are positive, and

$$\begin{aligned}&(\alpha-\phi)[\alpha(1-\rho_\theta\rho_z) + \rho_\theta(\rho_z-\alpha)] - (1-\alpha\phi)[(1-\rho_\theta\rho_z) - \rho_\theta(1-\alpha\rho_z)] \\ &\quad = (1-\alpha^2)[-1+\rho_\theta+\rho_\theta\rho_z(1-\phi)] < (1-\alpha^2)(\rho_\theta-\phi) < 0.\end{aligned}$$

Case 3. $H(L) = \frac{(1-\rho_\theta L)(1-\rho_z L)}{(L-\rho_\theta)(L-\rho_z)} A(L)$. Substituting this expression for $H(L)$ into the fixed point equation (83) and solving for $A(L)$ implies

$$A(L) = 1 - \frac{b_0 L[(1-\phi L)\psi(L) + (1-\alpha^2)\vartheta\sigma(L-\rho_\theta)^2(L-\rho_z)]}{(1-\lambda L)(1-\alpha L)(L-\alpha)(1-\phi L)}. \tag{93}$$

This equation and the hypothesis $A(\rho_\theta) = A(\rho_z) = 0$ implies that ϑ^2 satisfies (86) and (90). But it has been shown that (86) implies $\rho_\theta > \rho_z$ and (90) implies $\rho_\theta < \rho_z$, a contradiction.

Proof of Proposition 8. The proofs of the analogous versions of Propositions 1 and 4 are exactly the same as before, just with the relevant parameters indexed by i. The only thing that remains is to prove the analogous version of Proposition 2. By substituting (5) into (14), and averaging across sectors,

$$\bar{p}_t = \left[1 - \frac{1}{n}\sum_{i=1}^{n} \frac{b_{1i} f_{0i} \omega_i L}{(1-\lambda_i L)(1-\phi L)}\right] \bar{u}_t. \tag{94}$$

To prove the proposition, it is sufficient to verify that the operator on the right side is invertible into the past. This is true if and only if characteristic polynomial

$$\mathcal{P}(\mu) = \prod_{i=1}^{n}(\mu - \lambda_i)(\mu - \phi) - \frac{1}{n}\sum_{i=1}^{n} b_{1i} f_{0i}\omega_i \mu \prod_{j\neq i}(\mu - \lambda_j)$$

has $n+1$ inside zeros. Note first that $\operatorname{sign}\mathcal{P}(0) = (-1)^{n+1}$ because $0 < \phi < 1$ and $0 < \lambda_i < 1$, and $\mathcal{P}(1) > 0$ because

$$1 - \phi > \rho(1 - \phi/\rho)\frac{1}{n}\sum_{i=1}^{n} \frac{b_{1i} f_{0i}^2 (1-\beta_i\lambda_i)}{(b_{1i} f_{0i}^2 + f_{1i})(1-\beta_i\lambda_i\rho)} = \frac{1}{n}\sum_{i=1}^{n}\frac{b_{1i} f_{0i}\omega_i}{1-\lambda_i},$$

where the last equality uses the definitions of ω_i and λ_i.

Next, arrange the sequence $\{\lambda_i\}_{i=1}^{n}$ such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, and note that

$$\mathcal{P}(\lambda_i) = -\frac{1}{n}\sum_{i=1}^{n} b_{1i} f_{0i}\omega_i\lambda_i \prod_{j\neq i}(\lambda_i - \lambda_j).$$

This implies that $\mathcal{P}(\lambda_1) = 0$ if $\lambda_1 = \lambda_2$ and $\operatorname{sign}\mathcal{P}(\lambda_1) = (-1)^n$ otherwise, so $\operatorname{sign}\mathcal{P}(\lambda_1) \neq \operatorname{sign}\mathcal{P}(0)$. Similarly, $\operatorname{sign}\mathcal{P}(\lambda_n) = 0$ if $\lambda_n = \lambda_{n-1}$ and $\operatorname{sign}\mathcal{P}(\lambda_n) = -1$

otherwise, so $\operatorname{sign}\mathcal{P}(\lambda_n) \neq \operatorname{sign}\mathcal{P}(1)$. Finally, note that for $i = 2, 3, \ldots, n$,

$$\operatorname{sign}\mathcal{P}(\lambda_i) = \begin{cases} 0 & \lambda_i = \lambda_{i-1} \\ -\operatorname{sign}\mathcal{P}(\lambda_{i-1}) & \lambda_i > \lambda_{i-1} \end{cases}$$

Therefore $\mathcal{P}(\mu)$ has $n+1$ inside zeros.

Proof of Proposition 9. Existence and the uniqueness of the FCE follows from the same reasoning as in the proof of Proposition 1, and the policy function (6) is the same, except with the relevant structural parameters now indexed by i. Given (6), the closed-form expression in (1) comes from evaluating the forecasts $E_{it}u_{i,t+j}$ under the new law of motion (16). By(16) ,

$$E_{it}u_{i,t+j} = \alpha_i \rho^j E_{it}\theta_t \tag{95}$$

The vector $u_t = (u_{1t}, \ldots, u_{nt})$ contains the same information as $(\bar{u}_t, \hat{u}_{1t}, \ldots, \hat{u}_{2t})$, where $\bar{u}_t \equiv \frac{1}{n}\sum_{i=1}^n \frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i u_{it}$, $\sigma_\varepsilon^2 \equiv (\frac{1}{n}\sum_{i=1}^n \frac{\alpha_i^2}{\sigma_{\varepsilon i}^2})^{-1}$, and $\hat{u}_{it} \equiv u_{it} - \alpha_i \bar{u}_t$, since the two are related by a non-degenerate linear transformation. Since the processes $\{\hat{u}_{1t}, \ldots, \hat{u}_{nt}\}$ are independent of $\{\theta_t\}$,

$$E_{it}\theta_t = E(\theta_t|u^t) = E(\theta_t|\bar{u}^t, \hat{u}_1^t, \ldots, \hat{u}_n^t) = E(\theta_t|\bar{u}^t). \tag{96}$$

By (16), $\bar{u}_t = \theta_t + \sigma_\varepsilon \bar{\varepsilon}_t$, where $\bar{\varepsilon}_t \equiv \frac{1}{n}\sum_{i=1}^n \varepsilon_t$ is white noise with variance $1/n$. This is the same law of motion from the proof of Proposition 1, so the optimal forecast of θ_t is given by (27) (with appropriate re-definitions of ϕ and $\bar{u}_t$). Substituting (95), (96), and (27) into (6) delivers the same expression for $k_{i,t+1}$ presented in Proposition 1, with the new expression for ω_i reported in this proposition.

Substituting this policy function into the demand curve and computing the appropriately weighted average of prices across sectors implies

$$\bar{p}_t = \left[1 - \frac{1}{n}\sum_{i=1}^n \frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i^2 \frac{\tilde{\omega}_i L}{(1-\lambda_i L)(1-\phi L)}\right]\bar{u}_t, \tag{97}$$

where $\tilde{\omega}_i \equiv \omega_i/\alpha_i > 0$. To prove the analogous version of Proposition 2, it is sufficient to verify that the operator on the right side is invertible into the past. This is true if and only if characteristic polynomial

$$\mathcal{P}(\mu) = \prod_{i=1}^{n}(\mu - \lambda_i)(\mu - \phi) - \frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i^2\tilde{\omega}_i\mu\prod_{j\neq i}(\mu - \lambda_j)$$

has $n+1$ inside zeros. Note first that $\operatorname{sign}\mathcal{P}(0) = (-1)^{n+1}$ because $0 < \phi < 1$ and $0 < \lambda_i < 1$, and $\mathcal{P}(1) > 0$ because

$$1 - \phi > \rho(1 - \phi/\rho)\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i^2\frac{b_{1i}f_{0i}^2(1 - \beta_i\lambda_i)}{(b_{1i}f_{0i}^2 + f_{1i})(1 - \beta_i\lambda_i\rho)} = \frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i^2\frac{\tilde{\omega}_i}{1 - \lambda_i},$$

where the last equality uses the definitions of $\tilde{\omega}_i$ and λ_i.

Next, arrange the sequence $\{\lambda_i\}_{i=1}^n$ such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, and note that

$$\mathcal{P}(\lambda_i) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i^2\tilde{\omega}_i\lambda_i\prod_{j\neq i}(\lambda_i - \lambda_j).$$

This implies that $\mathcal{P}(\lambda_1) = 0$ if $\lambda_1 = \lambda_2$ and $\operatorname{sign}\mathcal{P}(\lambda_1) = (-1)^n$ otherwise, so $\operatorname{sign}\mathcal{P}(\lambda_1) \neq \operatorname{sign}\mathcal{P}(0)$. Similarly, $\operatorname{sign}\mathcal{P}(\lambda_n) = 0$ if $\lambda_n = \lambda_{n-1}$ and $\operatorname{sign}\mathcal{P}(\lambda_n) = -1$ otherwise, so $\operatorname{sign}\mathcal{P}(\lambda_n) \neq \operatorname{sign}\mathcal{P}(1)$. Finally, note that for $i = 2, 3, \ldots, n$,

$$\operatorname{sign}\mathcal{P}(\lambda_i) = \begin{cases} 0 & \lambda_i = \lambda_{i-1} \\ -\operatorname{sign}\mathcal{P}(\lambda_{i-1}) & \lambda_i > \lambda_{i-1} \end{cases}$$

Therefore, $\mathcal{P}(\mu)$ has $n+1$ inside zeros.

To prove the analogous version of Proposition 4, note that any REE of an economy with $s_{it} = \bar{p}_t$ implies a relationship of the form

$$\left[1 + \frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_\varepsilon^2}{\sigma_{\varepsilon i}^2}\alpha_i b_{1i}f_{0i}LA_i(L)\right]\bar{p}_t = \bar{u}_t.$$

The fact that $\bar{p}_t$ must be measurable with respect to $\xi^t = (v^t, \varepsilon^t)$ implies that the operator on the left must always be invertible into the past. This proves that, in any REE, $\bar{p}^t$ and $\bar{u}^t$ contain the same information.

B Numerical comparison to Han et al. (2022)

We solve the many sector Townsend model, parameterized as in Section 4.1, using our numerical procedure as well as that of Han et al. (2022). Han et al. (2022) provide general a purpose code for solving models with incomplete and dispersed information, something we have not done here. However, our procedure appears to provide some numerical advantages, at least for this particular model.

Our procedure requires specifying one approximation parameter, the density of the frequency grid in the interval $[-\pi, \pi]$ over which we approximate policy functions. For this, we consider values from 100 to 700 equally space points. The main approximation step in Han et al. (2022) requires specifying the orders p and q of the approximating ARMA(p, q) process used to solve the agents' inference problem. For this we consider values of $p = q$ from 1 to 7. We leave all other parameters of their algorithm at the default values set in their public code.

Figure 4 presents a time-precision frontier for the two algorithms. The vertical axis, in log scale, represents the time in second the procedure takes to complete. The horizontal axis captures the Euclidean norm of the residual of the aggregate impulse responses to the two aggregate shocks, computed over the first 50 periods. The residual is computed relative to a numerical solution using our algorithm and a 2500 point grid. The Figure shows that the frontier associated with out algorithm lies strictly to the southeast of Han et al. (2022), implying that this method always delivers higher accuracy for a given computation time.

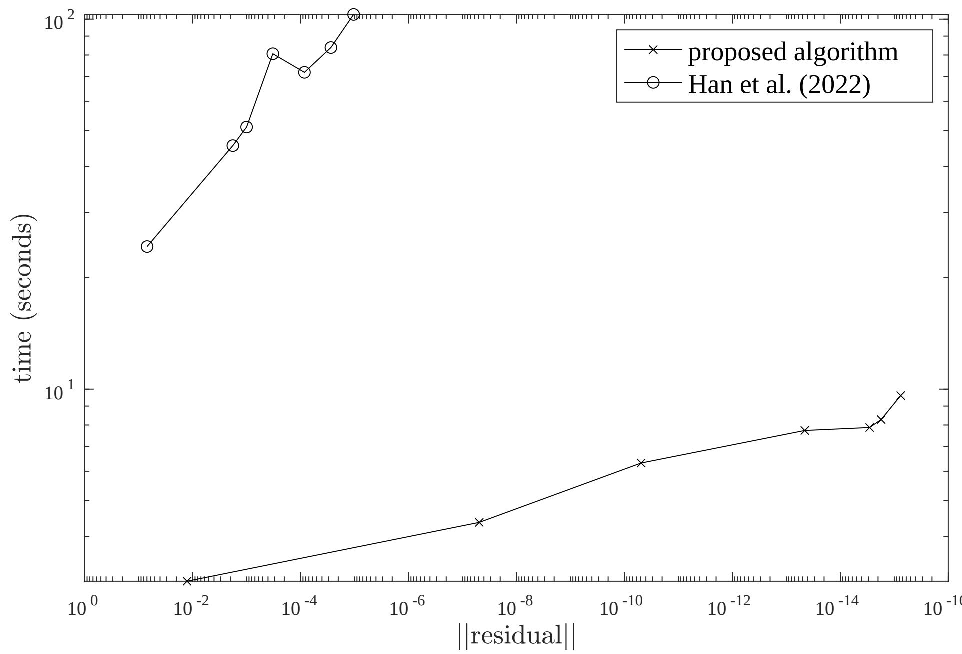

Figure 4: Comparison of numerical procedures.