KNOWING THE FORECASTS OF OTHERS

JOSEPH G. PEARLMAN AND THOMAS J. SARGENT

ABSTRACT. We apply recursive methods to obtain a finite dimensional and recursive representation of an equilibrium of one of Townsend's models of 'forecasting the forecasts of others'. The equilibrium has the property that decision makers make common forecasts of the hidden state variable whose presence motivates them to pay attention to prices in other markets. Thus, the model has too few sources of randomness to put decision makers into a situation where they should form 'higher order beliefs' (i.e., beliefs about others' beliefs). In Townsend's model, they know the beliefs of others because they share them. We attain our finite-dimensional recursive representation by applying methods of Pearlman, Currie, and Levine (1986).

Key Words: Forecasting the forecasts of others, higher order beliefs, pooling equilibrium, recursive methods, Kalman filter.

1. INTRODUCTION

Robert E. Lucas (1975), Kasa (2000), and Townsend (1983) make a compelling case that the assumption that decision makers have to extract signals about hidden persistent state variables is a good source both for additional impulses and for elongated impulse response functions in business cycle models. This theme has been pursued in recent analyses in which decision maker's imperfect information forces them into pursuing an infinite recursion of forming beliefs about the beliefs of other (e.g., Allen, Morris, and Shin (2002)). Robert E. Lucas (1975) side stepped the problem of forecasting the forecasts of others by letting decision makers pool their information before forecasting. Townsend (1983) bit the bullet, didn't assume pooling, and directly confronted the forecasting the forecasts of others problem. He proposed an approximate equilibrium of a model in which decision makers extract signals from endogenous variables (prices).

By applying results of Pearlman, Currie, and Levine (1986), this paper shows that there is a recursive representation of the equilibrium of the 'perpetually and symmetrically uninformed' model formulated but not completely solved in section 8 of Townsend (1983). Our computational method is recursive: it combines the Kalman filter with invariant subspace methods for solving systems of Euler equations.[1] As Singleton (1987), Kasa (2000), and Sargent (1991) also found, the equilibrium is fully revealing: observed prices tell participants in industry i all of the information held by participants in market $-i$ (i.e.,

Date: April 10, 2004.

The authors thank Kenneth Kasa for helpful comments on an earlier draft. Sargent gratefully acknowledges financial support from the National Science Foundation.

[1]See Anderson, Hansen, McGrattan, and Sargent (1996) for an account of invariant subspace methods.

'not i'). This means that higher-order beliefs play no role: seeing equilibrium prices lets decision makers pool their information sets.[2] The disappearance of higher order beliefs means that decision makers in this model do not really face a problem of forecasting the forecasts of others. They know those forecasts because they are the same as their own.

Townsend (1983) pointed out that his model with perpetually and symmetrically uninformed decision makers is one in which the state space explodes because it is necessary for decision makers to keep track of an infinite history of the vector of observables. He didn't actually characterize or compute an equilibrium for that model, but instead computed an approximate equilibrium by analyzing another model in which, after a finite number j periods, the lagged value of the key hidden state variable becomes known. The presence of that hidden state variable is the only thing that inspires decision makers in one market to condition their decisions on the history of prices in the other market. Sargent (1991) proposed a way to compute an equilibrium without making Townsend's approximation. Extending the reasoning of Muth (1960), Sargent noticed that it is possible to summarize the relevant history with a low dimensional object, namely, a small number of current and lagged forecasting errors. Positing an equilibrium in a space of perceived laws of motion for endogenous variables that takes the form of a vector autoregressive, moving average, Sargent described an equilibrium as a fixed point of a mapping from the perceived law of motion to the actual law of motion of that form. Sargent worked in the time domain and had to guess and verify the appropriate orders of the autoregressive and moving average pieces of the equilibrium representation. However, by working in the frequency domain Kasa (2000) showed how to discover the appropriate orders of the autoregressive and moving average parts, and also how to compute an equilibrium.

Our recursive computational method, which stays in the time domain, also discovers the appropriate orders of the autoregressive and moving average pieces. In addition, by displaying equilibrium representations in the form of Pearlman, Currie, and Levine (1986), we show how the moving average piece is linked to the innovation process of the hidden persistent component of the demand shock. That scalar innovation process is the additional state variable contributed by the problem of extracting a signal from equilibrium prices that decision makers face in Townsend's model.

2. Townsend's model

This section describes Townsend's basic model of an industry, then solves it under several assumptions about what decision makers observe. Firms in each of two industries $i = 1, 2$ employ a single factor of production capital, k_t^i, to produce output of a single good, y_t^i. We let capital letters denote market wide objects and lower case letters denote objects chosen by a representative firm. A representative firm in industry i has production function $y_t^i = f k_t^i$, $f > 0$, acts as a price taker with respect to output price P_t^i, and

[2] See Allen, Morris, and Shin (2002) for a discussion of the information assumptions needed to create a situation in which higher order beliefs appear in equilibrium decision rules. The way to read our findings in light of Allen, Morris, and Shin (2002) is that Townsend's section 8 model has too few sources of random shocks relative to sources of signals to permit higher order beliefs to play a role.

maximizes

$$E_0^i \sum_{t=0}^{\infty} \beta^t \left\{ P_t^i f k_t^i - .5h(k_{t+1}^i - k_t^i)^2 \right\}, \quad h > 0. \tag{2.1}$$

Demand in industry i obeys

$$P_t^i = -bY_t^i + \theta_t + \epsilon_t^i, \quad b > 0, \tag{2.2}$$

where $Y_t^i = fK_t^i$ is output in market i, θ_t is persistent component of a demand shock that is common across the two industries, and ϵ_t^i is an industry specific component of the demand shock that is i.i.d. with variance σ_ϵ^2. We assume that θ_t satisfies

$$\theta_{t+1} = \rho\theta_t + v_{t+1} \tag{2.3}$$

where $\{v_{t+1}\}$ is an i.i.d. sequence of Gaussian shocks with mean zero and variance σ_v^2. To simplify notation, we'll set $h = f = 1$. In equilibrium, $k_t^i = K_t^i$, but as usual we must distinguish between k_t^i and K_t^i when we pose the firm's optimization problem.

Townsend (1983) assumed that at time t firms in industry i observe $k_t^i, Y_t^i, P_t^i, (P^{-i})^t$, where $(P^{-i})^t$ is the history of prices in the other market up to time t. Notice that because the representative firm i sees the aggregate state variable Y_t^i in its own industry, as well as the price, it can infer the total demand shock $\theta_t + \epsilon_t^i$. However, at time t, the firm sees only P_t^{-i} and does not see Y_t^{-i}, so that firm i does not appear to see $\theta + \epsilon_t^{-i}$. Nevertheless, in the end the firm in industry i will be able to infer the composite shock $\theta_t + \epsilon_t^{-i}$ from the data that it does observe at t. We shall proceed to establish this result in steps.

2.0.1. *Strategy.* To prepare the way for solving Townsend's model, we shall first compute the law for capital in industry i under a sequence of assumptions about what the firm observes that make its information increasingly obscure. We begin with the most information, then gradually withdraw information in a way that approaches and eventually reaches the solution of Townsend's model that we are after. In particular, we shall consider the following information assumptions:

- Perfect foresight. Here we assume that future values of θ_t, ϵ_t^i are observed in industry i.
- Observed but stochastic θ_t. Here we assume that while $\{\theta_t, \epsilon_t^i\}$ are realizations from a stochastic process, current values of each are observed at time t.
- One noise-ridden observation on θ_t. At time t, a history w^t of a scalar noise-ridden observation on θ_t is observed at time t.
- Two noise-ridden observations on θ_t. At time t, a history w^t of *two* noise-ridden observations on θ_t is observed at time t.

Our solutions to these problems build one upon the other. We proceed by first finding the solution under perfect foresight. Then to get the solution with θ_t observed, we use a certainty equivalence principle to modify the perfect foresight solution by replacing future values of $\theta_s, \epsilon_s^i, s \geq t$ with expectations conditioned on θ_t. This gives us the solution when θ_t is observed. To get solutions when only the history of a noise ridden observation

w_t on θ_t is observed, we again apply a certainty equivalence principle and replace future values of $\theta_s, \epsilon_s^i, s \geq t$ with their expectations conditioned on w^t.

By this means, we shall discover solutions that form a benchmark against which we can interpret the equilibrium to Townsend's model that we shall compute in section 5 by applying the machinery of Pearlman, Currie, and Levine (1986). Our solution with two noise-ridden observations on θ_t will perfectly mimic the equilibrium of Townsend's model.

2.1. **Equilibrium conditions.** It is convenient to formulate the firm's problem as a discrete time Hamiltonian by forming the Lagrangian for the problem without uncertainty:

$$J = \sum_{t=0}^{\infty} \beta^t \left\{ P_t^i k_t^i - .5(\mu_t^i)^2 + \phi_t^i \left[k_t^i + \mu_t^i - k_{t+1}^i \right] \right\} \tag{2.4}$$

where $\{\phi_t^i\}$ is a sequence of Lagrange multipliers on the transition law for k_{t+1}^i. First order conditions for the nonstochastic problem are

$$\phi_t^i = \beta \phi_{t+1}^i + \beta P_{t+1}^i \tag{2.5}$$

$$\mu_t^i = \phi_t^i. \tag{2.6}$$

Substituting the demand function (2.2) for P_t^i, imposing the condition that the representative firm is representative ($k_t^i = K_t^i$), and using the definition below of g_t^i, the Euler equation (2.5), lagged by one period, can be expressed as $-bk_t^i + \theta_t + \epsilon_t^i + (k_{t+1}^i - k_t^i) - g_t^i = 0$ or

$$k_{t+1}^i = (b+1)k_t^i - \theta_t - \epsilon_t^i + g_t^i \tag{2.7}$$

where we define g_t^i by

$$g_t^i = \beta^{-1}(k_t^i - k_{t-1}^i). \tag{2.8}$$

We can write the Euler equation (2.5) in terms of g_t^i:

$$g_t^i = P_t^i + \beta g_{t+1}^i. \tag{2.9}$$

In addition, we have the law of motion for θ_t, (2.3), and the demand equation (2.2).

In summary, under the deterministic or perfect foresight interpretation, the equilibrium conditions for industry i consist of the following system of difference equations:

$$k_{t+1}^i = (1+b)k_t^i - \epsilon_t^i - \theta_t + g_t^i \tag{2.10}$$

$$\theta_{t+1} = \rho \theta_t + v_t \tag{2.11}$$

$$g_{t+1}^i = \beta^{-1}(g_t^i - P_t^i) \tag{2.12}$$

$$P_t^i = -bk_t^i + \epsilon_t^i + \theta_t \tag{2.13}$$

Without perfect foresight, the same system prevails except that the following equation replaces (2.12):

$$g_{t+1,t}^i = \beta^{-1}(g_t^i - P_t^i) \tag{2.14}$$

where $x_{t+1,t}$ denotes the expected value of x_{t+1} given information up to time t.

2.2. **Solution under perfect foresight.** Our first step is to compute the equilibrium law of motion for k_t^i under perfect foresight. Let L be the lag operator.[3] Equations (2.9) and (2.7) imply the second order difference equation in k_t^i:[4]

$$\left[(L^{-1}-(1+b))(1-\beta L^{-1})+b\right] k_t^i = \beta L^{-1}\epsilon_t^i + \beta L^{-1}\theta_t. \tag{2.15}$$

Factor the polynomial in L on the left side as:

$$-\beta[L^{-2}-(\beta^{-1}+(1+b))L^{-1}+\beta^{-1}] = \tilde{\lambda}^{-1}(L^{-1}-\tilde{\lambda})(1-\tilde{\lambda}\beta L^{-1}) \tag{2.16}$$

where $|\tilde{\lambda}|<1$ is the smaller root and λ is the larger root of $(\lambda-1)(\lambda-1/\beta)=b\lambda$. Therefore, (2.15) can be expressed as

$$\tilde{\lambda}^{-1}(L^{-1}-\tilde{\lambda})(1-\tilde{\lambda}\beta L^{-1})k_t^i = \beta L^{-1}\epsilon_t^i + \beta L^{-1}\theta_t. \tag{2.17}$$

Solving the stable root backwards and the unstable root forwards gives

$$k_{t+1}^i = \tilde{\lambda}k_t^i + \frac{\tilde{\lambda}\beta}{1-\tilde{\lambda}\beta L^{-1}}(\epsilon_{t+1}^i + \theta_{t+1}) \tag{2.18}$$

Thus under perfect foresight the capital stock satisfies

$$k_{t+1}^i = \tilde{\lambda}k_t^i + \sum_{j=1}^{\infty}(\tilde{\lambda}\beta)^j(\epsilon_{t+j}^i + \theta_{t+j}). \tag{2.19}$$

Next, we shall use alternative forecasting formulae in (2.19) to compute the equilibrium decision rule under alternative assumptions about the information available to decision makers in market i.

2.3. **Solution with θ_t stochastic but observed at t.** If future θ's are unknown at t, it is appropriate to replace all random variables on the right side of (2.19) with their conditional expectations based on the information available to decision makers in market i. In this section, we assume that this information set $I_t^p = \begin{bmatrix}\theta^t & \epsilon^{it}\end{bmatrix}$, where z^t represents the infinite history of variable z_s up to time t. In later subsections, we give the decision makers less and less information about θ_t.

To obtain the counterpart to (2.19) under our current assumption about information, we apply a certainty equivalence principle. In particular, it is legitimate to take (2.19) and replace each term $(\epsilon_{t+j}^i + \theta_{t+j})$ on the right side with $E[(\epsilon_{t+j}^i + \theta_{t+j})|\theta^t]$. After using (2.3) and the i.i.d. assumption about $\{\epsilon_t^i\}$, this gives

$$k_{t+1}^i = \tilde{\lambda}k_t^i + \frac{\tilde{\lambda}\beta\rho}{1-\tilde{\lambda}\beta\rho}\theta_t \tag{2.20}$$

or

$$k_{t+1}^i = \tilde{\lambda}k_t^i + \frac{\rho}{\lambda-\rho}\theta_t \tag{2.21}$$

[3] See Sargent (1987), especially chapters IX and XIV, for the methods used in this section.

[4] As noted by Sargent (1987), this difference equation is the Euler equation for the planning problem of maximizing the discounted sum of consumer plus producer surplus.

where $\lambda \equiv (\beta\tilde{\lambda})^{-1}$. For future purposes, it is useful to represent the solution for k_t^i recursively as

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{1}{\lambda - \rho}\hat{\theta}_{t+1} \tag{2.22}$$

$$\hat{\theta}_{t+1} = \rho\theta_t \tag{2.23}$$

$$\theta_{t+1} = \rho\theta_t + v_t. \tag{2.24}$$

3. Filtering

3.1. **One noisy signal.** We get closer to Townsend's setup, by now assuming that decision makers in market i do not observe θ_t, but that they do observe a history of noisy signals w^t. In particular, assume that

$$w_t = \theta_t + e_t \tag{3.25}$$

$$\theta_{t+1} = \rho\theta_t + v_t \tag{3.26}$$

where e_t and v_t are mutually independent i.i.d. Gaussian shock processes with means of zero and variances σ_e^2 and σ_v^2, respectively. Define

$$\hat{\theta}_{t+1} = E(\theta_{t+1}|w^t) \tag{3.27}$$

where w^t denotes the history of the w_s process up to and including t. Associated with the state-space representation (3.25),(3.26) is the *innovations representation*

$$\hat{\theta}_{t+1} = \rho\hat{\theta}_t + ka_t \tag{3.28}$$

$$w_t = \hat{\theta}_t + a_t \tag{3.29}$$

where $a_t \equiv w_t - E(w_t|w^{t-1})$ is the *innovations* process in w_t and the Kalman gain k is

$$k = \frac{\rho p}{p + \sigma_e^2} \tag{3.30}$$

and where p satisfies the Riccati equation

$$p = \sigma_v^2 + \rho^2 p - \frac{(\rho p)^2}{\sigma_e^2 + p}. \tag{3.31}$$

Define the state *reconstruction error* $\tilde{\theta}_t$ by

$$\tilde{\theta}_t = \theta_t - \hat{\theta}_t. \tag{3.32}$$

Then $p = E\tilde{\theta}_t\tilde{\theta}_t^T$. Equations (3.26) and (3.28) imply

$$\tilde{\theta}_{t+1} = (\rho - k)\tilde{\theta}_t + v_t - ke_t. \tag{3.33}$$

Now notice that we can express $\hat{\theta}_{t+1}$ as

$$\hat{\theta}_{t+1} = [\rho\theta_t + v_t] + [ke_t - (\rho - k)\tilde{\theta}_t - v_t], \tag{3.34}$$

where the first term in braces in the first line equals θ_{t+1} and the second term in braces equals $-\tilde{\theta}_{t+1}$.

3.1.1. *The θ-reconstruction error: a new state variable.* We can express (2.21) as

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{1}{\lambda - \rho} E\theta_{t+1}|\theta^t. \tag{3.35}$$

An application of a certainty equivalence principle asserts that when only w^t is observed, the appropriate solution is found by replacing the information set θ^t with w^t in (3.35). Making this substitution and using (3.34) leads to

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{\rho}{\lambda - \rho}\theta_t + \frac{k}{\lambda - \rho} e_t - \frac{\rho - k}{\lambda - \rho}\tilde{\theta}_t. \tag{3.36}$$

Simplifying equation (3.34), we also have

$$\hat{\theta}_{t+1} = \rho\theta_t + k e_t - (\rho - k)\tilde{\theta}_t. \tag{3.37}$$

Equations (3.36), (3.37) describe the solution when w^t is observed. Relative to (2.21), the solution acquires a new state variable, namely, the θ–reconstruction error, $\tilde{\theta}_t$. For future purposes, by using (3.30), it is useful to write (3.36) as

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{\rho}{\lambda - \rho}\theta_t + \frac{1}{\lambda - \rho}\frac{p\rho}{p + \sigma_e^2} e_t - \frac{1}{\lambda - \rho}\frac{\rho\sigma_e^2}{p + \sigma_e^2}\tilde{\theta}_t \tag{3.38}$$

In summary, when decision makers in market i observe one noisy signal on θ_t at t, we can write the equilibrium law of motion for k_t^i as

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{1}{\lambda - \rho}\hat{\theta}_{t+1} \tag{3.39}$$

$$\hat{\theta}_{t+1} = \rho\theta_t + \frac{\rho p}{p + \sigma_e^2} e_t - \frac{\rho\sigma_e^2}{p + \sigma_e^2}\tilde{\theta}_t \tag{3.40}$$

$$\tilde{\theta}_{t+1} = \frac{\rho\sigma_e^2}{p + \sigma_e^2}\tilde{\theta}_t - \frac{p\rho}{p + \sigma_e^2} e_t + v_t \tag{3.41}$$

$$\theta_{t+1} = \rho\theta_t + v_t. \tag{3.42}$$

3.2. **Two noisy signals.** We get even closer[5] to what will be the outcome in Townsend's model by assuming that the firm gets *two* noisy signals w_t on θ_t:

$$\theta_{t+1} = \rho\theta_t + v_t \tag{3.43}$$

$$w_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \theta_t + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \tag{3.44}$$

The innovations representation becomes

$$\hat{\theta}_{t+1} = \rho\hat{\theta}_t + k a_t \tag{3.45}$$

$$w_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \hat{\theta}_t + a_t \tag{3.46}$$

[5]'Even closer' is an understatement. As we shall see in section 5, we actually *match* the outcomes in the equilibrium of Townsend's model.

where $a_t \equiv w_t - E[w_t|w^{t-1}]$ is a (2×1) vector of innovations in w_t and k is now a (1×2) vector of Kalman gains. The formulae for the Kalman filter imply that

$$k = \frac{\rho p}{2p + \sigma_e^2} \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{3.47}$$

where $p = E\tilde{\theta}_t \tilde{\theta}_t^T$ now satisfies the Riccati equation

$$p = \sigma_v^2 + \frac{p\rho^2 \sigma_e^2}{2p + \sigma_e^2}. \tag{3.48}$$

In summary, when the representative firm in industry i observes *two* noisy signals on θ_t, we can express the equilibrium law of motion for capital recursively as

$$k_{t+1}^i = \tilde{\lambda} k_t^i + \frac{1}{\lambda - \rho} \hat{\theta}_{t+1} \tag{3.49}$$

$$\hat{\theta}_{t+1} = \rho \theta_t + \frac{\rho p}{2p + \sigma_e^2}(e_{1t} + e_{2t}) - \frac{\rho \sigma_e^2}{2p + \sigma_e^2} \tilde{\theta}_t \tag{3.50}$$

$$\tilde{\theta}_{t+1} = \frac{\rho \sigma_e^2}{2p + \sigma_e^2} \tilde{\theta}_t - \frac{p\rho}{2p + \sigma_e^2}(e_{1t} + e_{2t}) + v_t \tag{3.51}$$

$$\theta_{t+1} = \rho \theta_t + v_t. \tag{3.52}$$

We shall encounter versions of precisely these formulae again in section 5 where we compute the equilibrium of Townsend's model in which the representative firm in industry i receives a second noisy signal on θ_t by inferring it from P_t^{-i} and the other information that it has at time t. By extracting signals from the endogenous state variables, it will turn out that the firm recovers exactly the same process for the key additional state variable, the state reconstruction error $\tilde{\theta}_t$, that imperfect information contributes to the dynamics.

4. Method of PCL

In section 5, we shall compute the equilibrium of Townsend's model with two applications of the method of Pearlman, Currie, and Levine (1986). In this section, we briefly review a specialized version of their method that we shall apply. Readers who already know the Pearlman, Currie, and Levine procedures can proceed directly to section 5.

4.1. **Setup.** For any vector y_t, we let $y_{t+1,t} = E\left[y_{t+1}|I_t^p\right]$, where E is the mathematical expectation operator and I_t^p denotes the public's information set at time t, to be specified below. We shall make assumptions that imply that $[y_{t+1}, I_t^p]$ are jointly normally distributed, so that the mathematical expectation operator E coincides with the linear least squares projection operator. Let z_t denote a vector of *state* variables that are inherited from the past at t, while x_t is a vector of *jump* variables that adjust at time t to clear markets. We assume that decision makers' first-order necessary conditions and the other equilibrium conditions of our model can be arranged into the following special case of the setup of Pearlman, Currie, and Levine (1986):

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t} \end{bmatrix} = G \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} u_{1t} \\ 0 \end{bmatrix} \tag{4.53}$$

$$w_t = K \begin{bmatrix} z_t \\ x_t \end{bmatrix}. \tag{4.54}$$

Here u_{1t} is an i.i.d. Gaussian random vector with mean zero and covariance matrix $Eu_{1t}u_{1t}^T = U_{11}$ and w_t is a vector of variables whose history is observed by decision makers. That is, we assume that all decision makers have the information set

$$I_t^p = \{w_s : s \leq t\}. \tag{4.55}$$

In (4.53), (4.54), z_0 is a given initial condition, but x_0 is not given.

Let M be a matrix of right eigenvectors of G and let Λ be the diagonal matrix of the associated eigenvalues. Assume that the eigenvalues are distinct, so that we have the representation

$$MG = \Lambda M. \tag{4.56}$$

More generally, we can replace (4.56) with a Schur decomposition (see Anderson, Hansen, McGrattan, and Sargent (1996)). We assume that G has the saddle point property that the number of eigenvalues of G less than one in modulus is equal to the dimension of z_t and that the other eigenvalues are greater than one in modulus. Let Λ_2 contain the eigenvalues that exceed unity in modulus (call them the unstable eigenvalues) and let Λ_1 contain the eigenvalues that are less than unity in modulus (the stable eigenvalues), and partition M conformably with the partition of Λ. A *stable* solution of (4.53), (4.54) satisfies

$$E_0 \sum_{t=0}^{\infty} \gamma^t \begin{bmatrix} z_t \\ x_t \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix}^T < +\infty \tag{4.57}$$

for all $\gamma < 1$, given z_0. Under the saddle point property, there is a unique stable solution of our system that satisfies

$$M_{21} z_{t+1,t} + M_{22} x_{t+1,t} = 0. \tag{4.58}$$

4.2. **Solution.** We seek the stable solution of (4.53) that respects the information constraints that $x_{t+1,t}$ be measurable with respect to I_t^p. Following Pearlman, Currie, and Levine (1986), define the matrices[6]

$$\begin{aligned} C &= G_{11} - G_{12} M_{22}^{-1} M_{21} \\ A &= G_{11} - G_{12} G_{22}^{-1} G_{21} \\ D &= K_1 - K_2 G_{22}^{-1} G_{21} \end{aligned}$$

where G and K are conformably partitioned as

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \qquad K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$$

[6]We make the innocuous assumption that M_{22} is full rank; if not, then (4.58) implies that there is a linear combination of z_t that is unstable.

Let P be the unique positive definite matrix that satisfies the matrix Riccati equation

$$P = APA^T + U_{11} - APD^T \Delta DPA^T \tag{4.59}$$

where

$$\Delta = (DPD^T)^{-1}. \tag{4.60}$$

Define the innovation process $\tilde{z}_t = z_t - z_{t,t-1}$. The solution of the Riccati equation (4.59) determines the innovation covariance matrix $P = E\tilde{z}_t\tilde{z}_t^T$. Then the stable solution of (4.53) that respects the information constraints imposed on the expectations vector $x_{t+1,t}$ has the recursive representation

$$\tilde{z}_{t+1} = A(I - PD^T\Delta D)\tilde{z}_t + u_t^1 \tag{4.61}$$

$$z_{t+1} = Cz_t + (A - C)(I - PD^T\Delta D)\tilde{z}_t + u_t^1 \tag{4.62}$$

This solution incorporates two layers of optimization on the part of decision makers. First, decision makers' first-order conditions are arranged within (4.53), while the factorization (4.56) that is used in (4.59) solves stable roots backwards and unstable roots forwards, thereby assuring the stability of the solution. Second, decision makers' filtering and forecasting problems are solved via the Riccati equation (4.59).

5. Solving Townsend's model with PCL

Using the preceding results of Pearlman, Currie, and Levine (1986), we now compute the equilibrium of Townsend's model and verify that it mirrors the outcomes that we attained in the two-noisy-signal model of section 3.2. We proceed in two steps. First, we solve the model assuming common information or full information pooling.[7] That will generate a candidate decision rule that we shall use as a guess for the decision rule of firm $-i$. With that guess, we shall apply the methods of Pearlman, Currie, and Levine (1986) again to Townsend's setup with perpetually and symmetrically uninformed decision makers, and shall verify that the same decision rules prevail as in the pooling equilibrium. We shall see that in section 3 we have already encountered all of the key expressions that appear in those decision rules.

5.1. **Information pooling.** We temporarily assume that both industries are able to view values of k_t^i and P_t^i for both industries. This is precisely the system the equilibrium of which has already been determined in 3.2. Note that by augmenting the system with an additional state corresponding to each ϵ_t^i, it is possible to apply the methods of Pearlman, Currie, and Levine (1986), and reproduce equations (3.49) - (3.52). We consign the details to the appendix.[8]

[7] This is the type of equilibrium computed by Robert E. Lucas (1975).

[8] The appendix has the virtue of giving a cookbook example of how to apply the Pearlman, Currie, and Levine machinery.

5.2. **No information pooling.** As the key step of our 'guess-and-verify' strategy, we shall hypothesize that Corollary 3.1.1 of Kasa (2000) is true, namely that in equilibrium, the diverse information solution (agent i having information on P^i, P^j and k^i) leads to the same solution as in section 3.2. Thus from the point of view of industry 2, we shall assume that (3.49) - (3.52) hold for industry 1; we shall then prove that they hold for industry 2 as well.

We assume therefore that the system as viewed by industry 2 is given by

(5.63)
$$\begin{bmatrix} \varepsilon^1_{t+1} \\ \varepsilon^2_{t+1} \\ \theta_{t+1} \\ \widetilde{\theta}_{t+1} \\ k^1_{t+1} \\ k^2_{t+1} \\ g^2_{t+1,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 & 0 \\ \frac{-p\rho}{2p+\sigma^2_\varepsilon} & \frac{-p\rho}{2p+\sigma^2_\varepsilon} & 0 & \frac{\rho\sigma^2_\varepsilon}{2p+\sigma^2_\varepsilon)} & 0 & 0 & 0 \\ \frac{-p\rho}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \frac{-p\rho}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \frac{-\rho}{\rho-\lambda} & \frac{\rho\sigma^2_\varepsilon}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \tilde{\lambda} & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1+b & 1 \\ 0 & -1/\beta & -1/\beta & 0 & 0 & b/\beta & 1/\beta \end{bmatrix} \begin{bmatrix} \varepsilon^1_t \\ \varepsilon^2_t \\ \theta_t \\ \widetilde{\theta}_t \\ k^1_t \\ k^2_t \\ g^2_t \end{bmatrix} + \begin{bmatrix} \varepsilon^1_{t+1} \\ \varepsilon^2_{t+1} \\ v_t \\ v_t \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

with observations given by

(5.64)
$$w^2_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -b & 0 \\ 1 & 0 & 1 & 0 & -b & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon^1_t \\ \varepsilon^2_t \\ \theta_t \\ \widetilde{\theta}_t \\ k^1_t \\ k^2_t \\ g^2_t \end{bmatrix}$$

Here we have written the expression for k^1_{t+1} in terms of θ_t and $\widetilde{\theta}_t$ rather than $\widehat{\theta}_t$ Now the A, D and U matrices of section 4 are

(5.65)
$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 \\ \frac{-p\rho}{2p+\sigma^2_\varepsilon} & \frac{-p\rho}{2p+\sigma^2_\varepsilon} & 0 & \frac{\rho\sigma^2_\varepsilon}{2p+\sigma^2_\varepsilon)} & 0 & 0 \\ \frac{-p\rho}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \frac{-p\rho}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \frac{-\rho}{\rho-\lambda} & \frac{\rho\sigma^2_\varepsilon}{(2p+\sigma^2_\varepsilon)(\rho-\lambda)} & \tilde{\lambda} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(5.66)
$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & -b \\ 1 & 0 & 1 & 0 & -b & 0 \end{bmatrix} \qquad U = \begin{bmatrix} \sigma^2_\varepsilon & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_\varepsilon & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_v & \sigma^2_v & 0 & 0 \\ 0 & 0 & \sigma^2_v & \sigma^2_v & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Inspection of these matrices leads to the conclusion that the last state is deterministic and fully observable. Hence we can reduce the A and U matrices by removing the last row and column, with the D matrix changing to

$$D = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -b \end{bmatrix} \tag{5.67}$$

One can then show that the corresponding Riccati matrix is given by

$$P = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 & 0 & 0 \\ 0 & 0 & p & p & 0 \\ 0 & 0 & p & p & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.68}$$

where p is the same as equation (3.48). Hence

$$I - PD^T(DPD^T)^{-1}D = \begin{bmatrix} p & p & -\sigma_\varepsilon^2 & 0 & b(p+\sigma_\varepsilon^2) \\ p & p & -\sigma_\varepsilon^2 & 0 & -bp \\ -p & -p & \sigma_\varepsilon^2 & 0 & bp \\ -p & -p & -2p & \sigma_\varepsilon^2 + 2p & bp \\ 0 & 0 & 0 & 0 & \sigma_\varepsilon^2 + 2p \end{bmatrix} /(2p+\sigma_\varepsilon^2). \tag{5.69}$$

This time therefore, the equation for the innovations process corresponding to the first five states $\widetilde{z}_{t+1} = A(I - PD^T(DPD^T)^{-1}D)\widetilde{z}_t + u_t$ is represented by

$$\widetilde{z}_{t+1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{-p\rho}{2p+\sigma_\varepsilon^2} & \frac{-p\rho}{2p+\sigma_\varepsilon^2} & \frac{\rho\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2} & 0 & \frac{bp\rho}{2p+\sigma_\varepsilon^2} \\ \frac{-p\rho}{2p+\sigma_\varepsilon^2} & \frac{-p\rho}{2p+\sigma_\varepsilon^2} & 0 & \frac{\rho\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2} & 0 \\ 0 & 0 & \frac{-\rho\sigma_\varepsilon^2}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{\rho\sigma_\varepsilon^2}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{-p\rho b}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} \end{bmatrix} \widetilde{z}_t + \begin{bmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \\ v_t \\ v_t \\ 0 \end{bmatrix} \tag{5.70}$$

Note that $\widetilde{z}_{4t}$ satisfies an identical equation to $\widetilde{\theta}_t$. In addition, $\widetilde{z}_{5t}$ is dependent only on $\widetilde{z}_{3t}$-$\widetilde{z}_{4t}$, which itself depends on $\widetilde{z}_{5t}$, but not on any of the disturbance terms. Thus, both of these are zero, and it follows that $\widetilde{z}_{3t}$=$\widetilde{z}_{4t}$=$\widetilde{\theta}_t$.

The eigenvector of the unstable eigenvalue (which is the same as earlier) is given by [N 1] where $N = \begin{bmatrix} 0 & -1 & \frac{\lambda}{\rho-\lambda} & 0 & 0 & \lambda - \frac{1}{\beta} \end{bmatrix}$, so that

$$C = G_{11} - G_{12}N = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 \\ \frac{-p\rho}{2p+\sigma_\varepsilon^2} & \frac{-p\rho}{2p+\sigma_\varepsilon^2} & 0 & \frac{\rho\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2)} & 0 & 0 \\ \frac{-p\rho}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{-p\rho}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{-\rho}{\rho-\lambda} & \frac{\rho\sigma_\varepsilon^2}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \tilde{\lambda} & 0 \\ 0 & 0 & \frac{-\rho}{\rho-\lambda} & 0 & 0 & \tilde{\lambda} \end{bmatrix}. \tag{5.71}$$

Hence, after appending an additional row and column of zeros to $I - PD^T(DPD^T)^{-1}D$ as calculated above, we have

(5.72)

$$(A-C)(I-PD^T(DPD^T)^{-1}D) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-p\rho}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{-p\rho}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & \frac{\rho\sigma_\varepsilon^2}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & 0 & \frac{p\rho b}{(2p+\sigma_\varepsilon^2)(\rho-\lambda)} & 0 \end{bmatrix}$$

It therefore follows that $z_{t+1} = Cz_t + (A-C)(I-PD^T(DPD^T)^{-1}D)\widetilde{z}_t$ corresponds exactly to equations A.8, A.9, A.10, and A.11 in the appendix and that the dynamics are the same as for information pooling.

5.2.1. *Slight generalization.* So far we have assumed that both industries are identical, with $f = h = 1$. Suppose that adjustment costs differ for each industry. In particular, assume that $h = 1$ for industry 1, but not for industry 2. Then a similar approach to the above yields the same equation for k^1, while for industry 2,

$$k_{t+1}^2 = \tilde{\mu}k_t^2 - \frac{p}{2p+\sigma_\varepsilon^2}\frac{\rho}{\rho-\mu}(\varepsilon_t^1+\varepsilon_t^2) - \frac{\rho}{\rho-\mu}\theta_t + \frac{\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2}\frac{\rho}{\rho-\mu}\widetilde{\theta}_t \tag{5.73}$$

where μ and $\tilde{\mu}$ are the roots of $(\mu-1)(\mu-1/\beta) = b\mu/h$.

6. Concluding remarks

By applying recursive methods, this paper has verified a claim that has been suspected and that emerges in a sequence of papers including Townsend (1983), Singleton (1987), Sargent (1991), and Kasa (2000) in an environment like Townsend's. Townsend created that environment as a laboratory in which to study the effects of unleashing 'higher order beliefs'. He wanted to put traders 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 traders' private information. This result has both encouraging and discouraging aspects. Encouraging parts are that 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, that the resulting equilibria have low-dimensional representations, and that the signal extraction part of agents' problems give a new state variable – a state reconstruction error – that contributes interesting dynamics. A discouraging aspect is the fact that the dimension of the state-space *is* finite reflects the disappearance of the 'forecasting the forecasts of other problem' in equilibrium.

To reinstate the forecasting the forecasts of others problem, we would have to modify Townsend's environment to increase the dimension of decision makers' private information relative to the number of endogenous variables (prices) that decision makers can condition their forecasts upon. New work by Allen, Morris, and Shin (2002) shows that when this is done, the dimension of the state space can become infinite. One of the first macroeconomic applications of these ideas was by Woodford (2001), who used an islands model with a forecasting the forecasts of others feature to blow up the dimension of the state space and thereby get additional avenues for augmenting and shaping the persistence of the effects of monetary policy shocks on prices and output. In future work, we hope to apply the PCL technology that we have used here to that richer class of models. Preliminary research on an extension to Townsend's model by Sargent (1991) looks set to restore the infinite regress issue. The extension incorporates a cost of capital term in the firm's maximization problem, where the rental rate is stochastic. If the correlation between rental rates for the two firms is either 1 (i.e. impact of a shock is the same for each firm) or -1 (i.e. the aggregate shock is 0), then the solution is virtually identical to that obtained above, in that the Kalman filtering coefficients are the same. However if the correlation between rental rates lies between -1 and 1 then the system requires more than the two obvious additional state variables - each firm's estimate of the other firm's state estimation error. As yet, the authors have not found a finite state representation, or proved the necessity of infinite regress.

Appendix A. Application of PCL to pooling equilibrium

This appendix applies the methods of Pearlman, Currie, and Levine to verify the claim made in section 5.1. Assume that both industries are able to view values of k_t^i and P_t^i for both industries; the equations describing the system can be jointly written as:

$$\begin{bmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \\ \theta_{t+1} \\ k_{t+1}^1 \\ k_{t+1}^2 \\ g_{t+1,t}^1 \\ g_{t+1,t}^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 1+b & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1+b & 0 & 1 \\ -1/\beta & 0 & -1/\beta & b/\beta & 0 & 1/\beta & 0 \\ 0 & -1/\beta & -1/\beta & 0 & b/\beta & 0 & 1/\beta \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \theta_t \\ k_t^1 \\ k_t^2 \\ g_t^1 \\ g_t^2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \\ v_t \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.1}$$

with observations given by

$$\begin{bmatrix} k_t^1 \\ P_t^1 \\ k_t^2 \\ P_t^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -b & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -b & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \theta_t \\ k_t^1 \\ k_t^2 \\ g_t^1 \\ g_t^2 \end{bmatrix} \tag{A.2}$$

In the notation of Pearlman, Currie and Levine (1986) we can write this as

$$\begin{bmatrix} z_{t+1} \\ x_{t+1,t} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \end{bmatrix} \qquad w_t = \begin{bmatrix} K & 0 \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} \tag{A.3}$$

Now define the following matrices from PCL

$$(A.4)\qquad A=\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&\rho&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}\qquad D=\begin{bmatrix}0&0&0&1&0\\1&0&1&-b&0\\0&0&0&0&1\\0&1&1&0&-b\end{bmatrix}\qquad U_{11}=\operatorname{diag}(\sigma_\varepsilon^2,\sigma_\varepsilon^2,\sigma_v^2,0,0)$$

for which we must solve the Riccati equation $P = APA^T - APD^T(DPD^T)^{-1}DPA^T + U_{11}$. Inspection of these matrices leads to the conclusion that the 4th and 5th states are deterministic and fully observable, so that the innovations process for them is identically equal to zero. The matrices for the system with those states eliminated are then given by

$$(A.5)\qquad A=\begin{bmatrix}0&0&0\\0&0&0\\0&0&\rho\end{bmatrix}\qquad D=\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}\qquad U_{11}=\operatorname{diag}(\sigma_\varepsilon^2,\sigma_\varepsilon^2,\sigma_v^2).$$

This yields a solution to the Riccati equation given by $P=\operatorname{diag}(\sigma_\varepsilon^2,\sigma_\varepsilon^2,p)$ where

$$(A.6)\qquad p=\frac{\rho^2 p\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2}+\sigma_v^2.$$

After further effort we can show that for these three states

$$(A.7)\qquad I-PD^T(DPD^T)^{-1}D=\begin{bmatrix}p&p&-\sigma_\varepsilon^2\\p&p&-\sigma_\varepsilon^2\\-p&-p&\sigma_\varepsilon^2\end{bmatrix}/(2p+\sigma_\varepsilon^2)$$

and hence that the corresponding innovations process

$$(A.8)\qquad \widetilde{z}_{t+1}=A(I-PD^T(DPD^T)^{-1}D)\widetilde{z}_t+u_t$$

is given by

$$(A.9)\qquad \begin{bmatrix}\widetilde{\varepsilon}^1_{t+1}\\\widetilde{\varepsilon}^2_{t+1}\\\widetilde{\theta}_{t+1}\end{bmatrix}=\begin{bmatrix}0&0&0\\0&0&0\\\frac{-p\rho}{2p+\sigma_\varepsilon^2}&\frac{-p\rho}{2p+\sigma_\varepsilon^2}&\frac{\rho\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2}\end{bmatrix}\begin{bmatrix}\widetilde{\varepsilon}^1_{t}\\\widetilde{\varepsilon}^2_{t}\\\widetilde{\theta}_{t}\end{bmatrix}+\begin{bmatrix}\varepsilon^1_{t+1}\\\varepsilon^2_{t+1}\\v_t\end{bmatrix}.$$

Note that only the third state here is of interest, since the first two are identical to what we have already.

The equations for the main variables of the system are now given by

$$(A.10)\qquad z_{t+1}=Cz_t+(A-C)(I-PD^T(DPD^T)^{-1}D)\widetilde{z}_t+u_t$$

where

$$(A.11)\quad C=G_{11}-G_{12}N=\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&\rho&0&0\\0&0&\frac{-\rho}{\rho-\lambda}&\frac{1}{\beta\lambda}&0\\0&0&\frac{-\rho}{\rho-\lambda}&0&\frac{1}{\beta\lambda}\end{bmatrix}\qquad N=\begin{bmatrix}-1&0&\frac{\lambda}{\rho-\lambda}&\lambda-\frac{1}{\beta}&0\\0&-1&\frac{\lambda}{\rho-\lambda}&0&\lambda-\frac{1}{\beta}\end{bmatrix}$$

and λ is the larger solution of $(\lambda-1)(\lambda-1/\beta)=b\lambda$. (Note that $1/(\beta\lambda)$is therefore the smaller solution.) Also the matrix [N I] represents the two eigenvectors of the unstable eigenvalues of the G matrix.

To ensure that this is compatible with the 3-state innovations process derived above, we must append two rows of zeros to the matrix $I-PD^T(DPD^T)^{-1}D$.

The full representation of the reduced-form system can therefore be written as

$$k_{t+1}^i = \frac{1}{\beta\lambda} k_t^i - \frac{p}{2p+\sigma_\varepsilon^2}\frac{\rho}{\rho-\lambda}(\varepsilon_t^1 + \varepsilon_t^2) - \frac{\rho}{\rho-\lambda}\theta_t + \frac{\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2}\frac{\rho}{\rho-\lambda}\widetilde{\theta}_t \tag{A.12}$$

where

$$\widetilde{\theta}_{t+1} = \frac{\rho\sigma_\varepsilon^2}{2p+\sigma_\varepsilon^2}\widetilde{\theta}_t - \frac{p\rho}{2p+\sigma_\varepsilon^2}(\varepsilon_t^1 + \varepsilon_t^2) + v_t. \tag{A.13}$$

REFERENCES

Allen, F., S. Morris, and H. S. Shin (2002). Beauty contests, bubbles, and iterated expectations in asset markets. *mimeo*.

Anderson, E., L. P. Hansen, E. R. McGrattan, and T. J. Sargent (1996). Mechanics of forming and estimating dynamic linear economies. In H. M. Amman, D. A. Kendrick, and J. Rust (Eds.), *Handbook of computational economics*, pp. 171–252. Elsevier Science, North-Holland.

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

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association 55*, 299–306.

Pearlman, J., D. Currie, and P. Levine (1986). Rational Expectations Models with Private Information. *Economic Modelling 3*(2), 90–105.

Robert E. Lucas, J. (1975). An equilibrium model of the business cycle. *Journal of Political Economy 83*, 1113–1144.

Sargent, T. J. (1987). *Macroeconomic Theory, 2nd edition*. Academic Press.

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

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 Apprroaches to Monetary Economics*. Cambridge University Press.

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

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. *mimeo*.

JOSEPH G. PEARLMAN, DEPARTMENT OF ECONOMICS, LONDON METROPOLITAN UNIVERSITY
E-mail address: pearlman@lgu.ac.uk

THOMAS J. SARGENT, NEW YORK UNIVERSITY AND HOOVER INSTITUTION, STANFORD, CALIFORNIA
E-mail address: thomas.sargent@nyu.edu