

Simple Models and Biased Forecasts^{*}

Pooya Molavi[†]

September 12, 2022

This paper proposes a framework in which agents are constrained to use simple time-series models to forecast economic variables and characterizes the resulting biases. It considers agents who can only entertain state-space models with no more than d states, where d measures the agents' cognitive abilities. When the true data-generating process does not have a d -state representation, agents end up with misspecified models and biased forecasts. Under some assumptions, agents attend to the most persistent observables at the expense of less persistent ones. This bias anchors agents' forward-looking decisions to persistent state variables and increases comovement among those decisions. The paper proceeds to study the implications of the theory in the context of new-Keynesian, real business cycle, and Diamond–Mortensen–Pissarides models. In each case, constraining agents to use simple models brings the outcomes more in line with stylized facts.

^{*}This paper subsumes and substantially extends the second half of my job-market paper, titled “Macroeconomics with Learning and Misspecification: A General Theory and Applications.” I am grateful to Marios Angeletos, Simon Board, Larry Christiano, Nico Crouzet, Ian Dew-Becker, Théo Durandard, Marty Eichenbaum, Jeff Ely, Timothy Feddersen, Laura Gáti, Yuriy Gorodnichenko, Lars Hansen, Ali Jadbabaie, Chen Lian, Annie Liang, Joshua Mollner, Emi Nakamura, Kris Nimark, Serena Ng, Alessandro Pavan, Giorgio Primiceri, Matt Rognlie, Andy Skrzypacz, Jón Steinsson, Bruno Strulovici, Alireza Tahbaz-Salehi, Harald Uhlig, Vish Viswanathan, Olivier Wang, Iván Werning, Alex Wolitzky, Mike Woodford, and seminar participants at Cornell University, Northwestern University, UC Berkeley, Saïeh Family Fellows Conference, Brown University, Northwestern-UCChicago Economic Theory Conference, Pizzanomics Conference (EIEF), NBER Summer Institute, SITE, and Econometric Society European Meeting for discussions and their helpful comments.

[†]Northwestern University, pmolavi@kellogg.northwestern.edu.

1 Introduction

When faced with the difficult task of forecasting in a complex world, people are bound to rely on simple models and past experiences. Yet, the rational expectations hypothesis submits that agents can forecast the future as if they knew the true model of the economy. The learning literature (e.g., [Marcet and Sargent \(1989\)](#), [Evans and Honkapohja \(1995\)](#)) addresses this criticism of rational expectations by making agents like econometricians who estimate the model of the economy using available data. This paper proposes a bounded rationality model that builds on this insight.

I propose a framework in which agents use flexible but simple time-series models that are fit to data to make sense of their environment. An agent attempts to forecast future values of a set of observables based on their past realizations. The observables follow a stochastic process that is unknown to the agent. Instead, she starts with the prior that any model in the flexible class of state-space models of dimension d is plausible, and she uses her observations and Bayes' rule to form her forecasts. If the true process does not have a state-space representation of dimension d , the simple models she considers will be misspecified. Consequently, the agent will not learn the true model, even in the long run. Her belief will instead asymptotically concentrate on the set of *pseudo-true models*, which provide low-dimensional approximations of the true process.

The paper's main contribution is characterizing the biases that result from agents' use of pseudo-true state-space models. I begin my analysis by establishing a useful linear-invariance property of pseudo-true models: Expectations of agents who use pseudo-true state-space models respect all linear *intratemporal* relationships among observables; it is only *intertemporal* statistical relationships that may be misperceived by such agents. The result implies that the way agents form their forecasts is invariant to the scale or volatility of different observables. The linear-invariance property makes the framework immediately applicable in macro applications, even when there is no unique or obvious way of defining the vector of observables.

The paper's main characterization result establishes that (under appropriate assumptions) the use of pseudo-true state-space models leads to *persistence bias*: Agents forecast the most persistent components of the vector of observables as well as possible while missing the dynamics of the remaining components. Furthermore, there is generally a unique way of ranking the persistence of observables: Different agents—with different payoffs, facing different decisions, and using pseudo-true models with possibly different dimensions—all agree on what the most persistent components of the observables are.

Persistence bias has further testable implications. The forecasts and forward-looking actions of agents who use pseudo-true state-space models are unresponsive to innovations in the least persistent observables. They are instead anchored to sluggish, backward-looking observables. Persistence bias also leads agents' forward-looking decisions to comove with each other. Actions comove regardless of the specifics of the agents' decision problems or the dimension of

their models, as long as actions are all sufficiently forward-looking and agents use pseudo-true models of sufficiently low dimension.

The framework of simple models provides a parsimonious and tractable way of introducing bounded rationality into standard macro models. The only free parameter in the specification of agents' expectations is dimension d of their models. If d is sufficiently large, the agents' expectations revert to rational expectations. Meanwhile, other parameters of agents' subjective models are determined endogenously as they fit their models to the true process.

One can easily incorporate the framework in large-scale macroeconomic models. The set of pseudo-true state-space models is independent of the agent's preferences. The problem of finding pseudo-true models thus decouples from the agent's optimal decision problem. Furthermore, the characterization results of the paper give close-form expressions (under some assumptions) for the set of pseudo-true state-space models as functions of d and the second moments of the true process.

The paper's theoretical results offer a powerful toolbox to study how bounded rationality alters the predictions of standard macroeconomic models and the associated policy prescriptions. I illustrate these tools in the context of three workhorse macro models: the new-Keynesian model, the real business cycle (RBC) model, and the Diamond–Mortensen–Pissarides (DMP) model. In each case, constraining agents to have simple models brings the predictions more in line with stylized facts.

As the first application, I consider a version of the standard new-Keynesian model in which agents are constrained to forecast using pseudo-true one-dimensional models. As in the rational expectations version of the model, the equilibrium has a simple linear representation, which can be computed analytically. I then use the equilibrium characterization to study the implications of bounded rationality for the conduct of monetary policy.

Several insights arise from the analysis of the new-Keynesian model. First, even if agents employ simple models, the monetary authority can simultaneously achieve zero output gap and inflation (as long as cost-push shocks are identically zero). Second, the agents' reliance on simple statistical models that are fit to past data curtails the power of forward guidance. Moreover, the power of forward guidance is largely independent of the guidance duration.

The second application concerns the propagation of productivity shocks in the RBC model. The RBC model is an excellent case study because it has only one exogenous shock and two state variables. Therefore, if agents are constrained to d -dimensional models with $d \geq 2$, they recover the true process, and their expectations coincide with rational expectations. When $d = 1$, on the other hand, agents' models will be pseudo-true, and their forecasts will be biased. This prediction of the framework of simple models distinguishes it from signal-extraction-type models, which revert to rational expectations when there is a single exogenous shock in the economy.

Constraining agents to one-dimensional models in the RBC model causes aggregate consumption to behave more like a stock variable. This prediction directly follows from the persis-

tence bias in agents' expectations. Agents' estimate of the subjective state mostly depends on the value of the most persistent variable—the capital stock in equilibrium for the RBC model. Consequently, consumption (an almost purely forward-looking variable) comoves with the value of the capital stock. The anchoring of consumption to the capital stock makes consumption more sluggish than under rational expectations and increases the volatilities of consumption, hours, and investment.

As the last application, I study how the predictions of the standard search and matching model change when agents are constrained to use simple models. I consider a standard calibration of the DMP model with labor productivity and separation rate shocks. The equilibrium has a three-state representation, so expectations of agents constrained to d -dimensional models with $d \geq 3$ coincide with rational expectations. I instead consider agents constrained to use one-dimensional models. In equilibrium, agents' estimate of the subjective state closely tracks the evolution of the unemployment rate. Separation rate shocks increase the unemployment rate, thus making agents pessimistic about the state of the economy. The result is a decrease in vacancy creation following an increase in separations and a negative comovement between the unemployment rate and vacancies in response to the separation shock. Meanwhile, the stickiness of expectations slows the dynamics of the economy, thus improving the propagation mechanism of the model.

Related Literature. This paper belongs to the growing literature on models and implications of deviating from the benchmark of full-information rational-expectations in macroeconomics. See [Woodford \(2013\)](#) for a survey.

Within this literature, this paper is closest in spirit to the adaptive learning literature, which goes back to [Marcet and Sargent \(1989\)](#), [Sargent \(1993, 1999\)](#), and [Evans and Honkapohja \(1995\)](#).¹ The paper shares this literature's approach to modeling economic agents as econometricians who use estimated statistical models to make sense of their world. However, it deviates from the earlier works in that literature in several important ways. First, I assume that the agents' model of the economy is a state-space model, which is fully flexible except in its dimension. Second, I consider Bayesian agents and focus on the limit where learning is complete and agents have settled on some pseudo-true model. Finally, I accommodate the possibility that the agents' pseudo-true models differ from the true model due to misspecification.

An alternative way of deviating from the benchmark model is to assume imperfect knowledge of payoff-relevant variables due to either dispersed information, e.g., [Lucas \(1972\)](#), noisy information, e.g., [Orphanides \(2003\)](#) and [Angeletos and La'O \(2009\)](#), sticky information, e.g., [Mankiw and Reis \(2002\)](#), or costly attention, e.g., [Sims \(2003\)](#), [Woodford \(2003\)](#), [Maćkowiak and Wieder-](#)

¹Other related papers in the learning literature include [Cho, Williams, and Sargent \(2002\)](#), [Bullard and Mitra \(2002\)](#), [Marcet and Nicolini \(2003\)](#), [Orphanides and Williams \(2005\)](#), [Preston \(2005\)](#), [Milani \(2007\)](#), [Bullard, Evans, and Honkapohja \(2008\)](#), [Adam and Marcet \(2011\)](#); [Adam, Marcet, and Beutel \(2017\)](#), [Eusepi and Preston \(2011\)](#), [Gerko \(2017\)](#), [Gáti \(2020\)](#), and [Kozłowski, Veldkamp, and Venkateswaran \(2020\)](#). See [Evans and Honkapohja \(2012\)](#) for a textbook treatment and [Eusepi and Preston \(2018\)](#) for a recent literature survey.

holt (2009, 2015), and Gabaix (2014).² This paper abstracts from the difficulty of observing a large cross-section of variables and instead focuses on the difficulty of comprehending complex time-series relationships. The predictions of this framework also distinguish it from the literature mentioned above: In my model, agents fully uncover cross-sectional relationships among variables, but their expectations could deviate from rational expectations even if the economy has a single exogenous shock.

This paper also contributes to the literature that studies the properties of pseudo-true models. The term pseudo-true model originates in the pioneering work of Sawa (1978), who proposes using KL divergence as a model-selection criterion when models are misspecified. Agents in the restricted-perceptions equilibrium of Bray (1982) and Bray and Savin (1986), Rabin and Vayanos (2010)’s model of the gambler’s fallacy, the natural-expectations framework of Fuster, Laibson, and Mendel (2010), and Fuster, Hebert, and Laibson (2012), the Berk–Nash equilibrium of Esponda and Pouzo (2016, 2021), and the constrained rational expectations equilibrium of Molavi (2019) all use pseudo-true models to forecast payoff-relevant variables. Agents in Krusell and Smith (1998) also have a misspecified model of the economy. They believe that current and future prices do not depend on anything but the first few moments of the wealth distribution.

However, despite this long history, surprisingly few general results on the properties of pseudo-true models have appeared in the literature. Such results are almost exclusively derived—with the notable exception of Rabin and Vayanos (2010)—in settings where the set of models is sufficiently restricted that the pseudo-true model can be estimated using OLS regression, and the bias in agents’ forecasts reduces to the omitted-variable bias. I contribute to this literature by characterizing the set of pseudo-true state-space models of a given dimension.

A large empirical literature studies the question of whether households, firms, and professional forecasters underreact or overreact to new information. Coibion and Gorodnichenko (2015) provide evidence of underreaction in consensus forecasts for professional forecasters. Bordalo, Gennaioli, Ma, and Shleifer (2020) show that individual forecasts of professional forecasters overreact to news. Broer and Kohlhas (2020) find evidence for both overreaction and underreaction depending on the aggregate variable being studied. More recently, Angeletos, Huo, and Sastry (2021) find evidence of underreaction at short horizons and overreaction at longer horizons, whereas Afrouzi, Kwon, Landier, Ma, and Thesmar (2021) find evidence of overreaction in a lab setting.

In parallel with this empirical literature, many authors have proposed theoretical models of overreaction and underreaction. Natural expectations, e.g., Fuster, Laibson, and Mendel (2010) and Fuster, Hebert, and Laibson (2012) and diagnostic expectations, e.g., Bordalo, Gennaioli, and Shleifer (2018) and Bianchi, Ilut, and Saijo (2021) are examples of models where agents overextrapolate from the recent past. Cognitive discounting of Gabaix (2020) and level- k thinking, e.g., García-Schmidt and Woodford (2019) and Farhi and Werning (2019) are examples of models

²See also Nimark (2008), Lorenzoni (2009), Alvarez, Lippi, and Paciello (2015), Angeletos and Lian (2018), Angeletos and Huo (2021), and Chahrour, Nimark, and Pitschner (2021).

that feature underreaction. The framework proposed in this paper is neither a model of underreaction nor of overreaction. In fact, I show that agents who use simple models always overreact to some observables and underreact to others.

The state-space models used in this paper are relatives of dynamic factor models, e.g., [Stock and Watson \(2011, 2016\)](#). However, the two are distinct both mathematically and conceptually. Dynamic factor and state-space models offer two different representations of stochastic processes.³ Each representation suggests a conceptually different decomposition of time-series data. Dynamic factor models decompose data into common factors and idiosyncratic disturbances, whereas state-space models decompose it into persistent and transitory components. The two approaches thus suggest two different simplifications of large time-series data: using a small number of common factors in the former case and a small number of persistent states in the latter.

Finally, [Molavi, Tahbaz-Salehi, and Vedolin \(2021\)](#) use a closely related framework to study the implications of model misspecification for asset prices and returns. They show that constraining the complexity of investors' models leads to return predictability and provides a parsimonious account of several puzzles in the asset-pricing literature.

Outline. The rest of the paper is organized as follows: Section 2 presents the framework of simple models and formally defines and discusses the notion of a pseudo-true model. Section 3 contains the paper's characterization results for the set of pseudo-true state-space models. Section 4 discusses the implications of using pseudo-true simple models for agents' forecasts and choices. Section 5 discusses an application to monetary policy in the new-Keynesian model. Section 6 shows how constraining agents to use simple models alters the amplification and propagation of productivity shocks in the RBC model. Section 7 contains the labor search and matching application. Section 8 concludes. Some additional results are provided in two appendices. The proofs of theoretical results and other calculations are relegated to the online appendices.

2 General Framework

In this section, I present the environment and the main behavioral assumption of the paper.

2.1 The Environment

Time is discrete and is indexed by $t \in \mathbb{Z}$. An agent observes a sequence of variables over time and uses her past observations to forecast their future values. I let $y_t \in \mathbb{R}^n$ denote the time- t value of the vector of observables, or simply the *observable*. Vector y_t follows a mean-zero stochastic process \mathbb{P} with the corresponding expectation operator $\mathbb{E}[\cdot]$. I start by taking \mathbb{P} as a primitive; the

³The sets of time series that can be represented by dynamic factor and state-space models are not nested. Instead, any finite dynamic factor model has a state-space representation, and any finite state-space model has a dynamic factor representation. See [Forni and Lippi \(2001\)](#) for a representation result for the (generalized) dynamic factor models.

process will be an endogenous outcome of agents' actions in the macro applications studied in Sections 5-7.

I make several technical assumptions on the true process. First, \mathbb{P} is stationary ergodic with $\mathbb{E}[\|y_t\|^2] < \infty$. Second, it is purely non-deterministic. Third, there exists a linear subspace \mathcal{W} of \mathbb{R}^n (possibly equal to \mathbb{R}^n itself) such that y_t is supported on \mathcal{W} . Third, $\mathbb{P}(y_1, \dots, y_t)$ is absolutely continuous with respect to the Lebesgue measure on \mathcal{W}^t for any t , with density $\mathbb{f}(y_1, \dots, y_t)$.⁴ Finally, the true process has finite entropy rate, i.e., $\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}[-\log \mathbb{f}(y_1, \dots, y_t)] < \infty$. These assumptions are all quite weak. They are satisfied, for instance, if y_t follows a mean-zero vector ARMA process with Gaussian innovations.

The agent has perfect information about the past realizations of the observable, with her time- t information set given by $\{y_t, y_{t-1}, \dots\}$. However, she may use a misspecified model to map her information to her forecasts. This model misspecification leads to deviations in the agent's forecasts from those that arise in the rational-expectations benchmark.

2.2 Simple Models

As the paper's main behavioral assumption, I assume that the agent is constrained to use state-space models with a small number of state variables to forecast the vector of observables. She can only entertain models of the form

$$\begin{aligned} z_t &= Az_{t-1} + w_t, \\ y_t &= B'z_t + v_t, \end{aligned} \tag{1}$$

where z_t is the d -dimensional vector of *subjective latent states*, $A \in \mathbb{R}^{d \times d}$, $w_t \in \mathbb{R}^d$ is i.i.d. $\mathcal{N}(0, Q)$, $B \in \mathbb{R}^{d \times n}$, $v_t \in \mathbb{R}^n$ is i.i.d. $\mathcal{N}(0, R)$, and w_t and v_t are independent. While the integer d is a primitive of the model that parameterizes the dimension of the agent's models, matrices A , B , Q , and R are parameters that are determined endogenously by maximizing the fit to the true process.

Formally, I define a *d-state model* as a stationary ergodic stochastic process for y_t that has a representation of the form (1) such that (i) the dimension of vector z_t is d , (ii) A is a convergent matrix, (iii) Q is positive definite, and (iv) R is positive semidefinite.⁵ Whenever there is no risk of confusion, I use the term *d-state model* to refer both to the stochastic process for y_t and the parameters $\theta \equiv (A, B, Q, R)$ of its state-space representation. I let Θ_d denote the set of all d -state models, let P^θ denote the stationary distribution over $\{y_t\}_{t=-\infty}^\infty$ induced by model θ , and let $\mathcal{P}_d \equiv \{P^\theta : \theta \in \Theta_d\}$.⁶ With slight abuse of notation, I write $\Theta_d \subseteq \Theta_{d+1}$ to stress the fact that agents with larger values of d can entertain larger classes of models.

⁴This assumption is weaker than the assumption that $\mathbb{P}(y_1, \dots, y_n)$ is absolutely continuous with respect to the Lebesgue measure over \mathbb{R}^{nt} since it allows for the possibility that the true process is degenerate. This additional level of generality will be useful in applications where the elements of y_t may be linearly dependent.

⁵A matrix is *convergent* if all of its eigenvalues are smaller than one in magnitude. A being convergent and Q being positive definite are sufficient for a model (A, B, Q, R) to define a purely non-deterministic stationary ergodic process.

⁶One can define the set of d -state models without any reference to the latent state z_t . Stochastic process P for $\{y_t\}_{t=-\infty}^\infty$ with expectation operator E is a d -state model if $E[y_t y_{t-l}'] = CA^{l-1}\bar{C}'$ for all $l = 1, 2, \dots$, some convergent $d \times d$ matrix A , and some $C, \bar{C} \in \mathbb{R}^{n \times d}$. See, for instance, Faure (1976) or Katayama (2005, Chapter 7). I opt for the definition that uses the subjective latent state since z_t will have an intuitive interpretation as the state of the economy in the macro applications I consider in this paper.

The integer d captures the agent’s sophistication in modeling the stochastic process for the vector of observables, with larger values of d amounting to agents who can entertain more complex models. When d is sufficiently large, by the Wold representation theorem, the agent’s set of models contains arbitrarily good approximations to any purely non-deterministic covariance-stationary process. On the other hand, when d is small relative to the number of states required to model the true process, no model in the agent’s set of models will provide a good approximation to \mathbb{P} . The agent then necessarily ends up with a misspecified model of the true process and biased forecasts—regardless of which model in the set \mathcal{P}_d she uses to make her forecasts. Characterizing this bias is the focus of the next section of the paper.

My preferred rationale for the constraint on the number of states is to capture limits on the agent’s cognitive abilities, but the constraint can also arise from the agent’s rational fear of overfitting. Models with a large number of parameters and many degrees of freedom are prone to overfitting. Such concerns may lead rational agents to limit themselves to statistical models with a small number of parameters, especially if they only have a short time series to draw upon when estimating the parameters of their model. In the remainder of the paper, I abstract away from any issues arising from small samples and instead consider the long-run limit where the sampling error vanishes.

2.3 Pseudo-True Models

I assume that the agent forecasts using a model in the family of d -state models that provides the best fit to the true process. I use the Kullback–Leibler divergence rate of process P^θ from the true process \mathbb{P} as the measure of the fit of model θ .⁷ The *Kullback–Leibler divergence rate* (KLDR) of P^θ from \mathbb{P} is denoted by $\text{KLDR}(\theta)$ and defined as follows. Recall that the true process is supported on a linear subspace \mathcal{W} of \mathbb{R}^n . If P^θ is also supported on \mathcal{W} , then

$$\text{KLDR}(\theta) \equiv \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\log \left(\frac{f(y_1, \dots, y_t)}{f^\theta(y_1, \dots, y_t)} \right) \right],$$

where $f^\theta(y_1, \dots, y_t)$ denotes the density of P^θ with respect to the Lebesgue measure on \mathcal{W}^t ; if P^θ is not supported on \mathcal{W} , then $\text{KLDR}(\theta) \equiv +\infty$.

The Kullback–Leibler divergence rate is the natural generalization of Kullback–Leibler (KL) divergence to stationary stochastic processes. In the i.i.d. case, the KL divergence of a candidate model from the true model captures the difficulty of rejecting the candidate model in favor of the true model using a likelihood-ratio test. That is why the KL divergence is commonly used as a measure of a model’s fit.⁸ Similarly, $\text{KLDR}(\theta)$ captures the rate at which the power of a test for separating a stochastic process P^θ from the true process \mathbb{P} approaches one as $t \rightarrow \infty$.⁹ The KLDR is also tightly linked to asymptotics of Bayesian learning, as I discuss in the following subsection.

⁷The mean-squared forecast error is another commonly used notion of fit. In Appendix A, I define the weighted mean-squared forecast error and show that it is equivalent to the Kullback–Leibler divergence rate under an appropriate choice of the weighting matrix.

⁸See, for instance, Hansen and Sargent (2008).

⁹See, for instance, Shalizi (2009).

Model $\theta^* \in \Theta_d$ is a *pseudo-true d -state model* if $\text{KLDR}(\theta^*) \leq \text{KLDR}(\theta)$ for all $\theta \in \Theta_d$. If the agent’s set of models contains a model θ such that $f^\theta(y_1, \dots, y_t) = \mathbb{f}(y_1, \dots, y_t)$ almost everywhere and for all t , then any pseudo-true d -state model is observationally equivalent to the true process.¹⁰ The set \mathcal{P}_d of distributions is then correctly specified. When no such d -state model exists, $\text{KLDR}(\theta) > 0$ for any model $\theta \in \Theta_d$, and the set \mathcal{P}_d is misspecified. I let Θ_d^* denote the set of all pseudo-true d -state models, and let $\mathcal{P}_d^* \equiv \{P^\theta : \theta \in \Theta_d^*\}$. The following result shows that pseudo-true models are observationally equivalent to the true process when the set of models is correctly specified:

Theorem 1. *Suppose the set \mathcal{P}_d of d -state models is correctly specified. Then any pseudo-true d -state model $P^{d*} \in \mathcal{P}_d^*$ is observationally equivalent to the true process \mathbb{P} .*

The paper’s focus is the misspecified case, where d is small relative to the number of states required to capture the true process. This is a statement about d being smaller than the “true d ”—and not it being smaller than n , the dimension of y_t . However, it is often natural to also think of d as much smaller than n . Approximating the true process by a pseudo-true d -state model then amounts to using a parsimonious time-series model to capture the essential features of a large data set. Unless otherwise specified, I assume throughout the paper that $d \leq n$, but the paper’s characterization results easily generalize to the $d > n$ case.

2.4 Learning Foundation

Pseudo-true models arise naturally as the long-run outcome of learning by Bayesian agents with misspecified priors. Consider an agent who starts with prior μ_0 with full support over the points in the set $\mathbb{R}^d \times \Theta_d$, each corresponding to an initial value of the subjective states z_0 and a d -state model θ , which describes how states and the observable coevolve. Suppose the agent observes y_t over time and updates her belief using Bayes’ rule. Let μ_t denote the agent’s time- t Bayesian posterior over $\mathbb{R}^d \times \Theta_d$. Berk (1966)’s theorem establishes that, in the limit $t \rightarrow \infty$, the agent’s posterior will assign a probability of one to the set of pseudo-true models.¹¹

This result offers an “as if” interpretation of pseudo-true d -state models. One can assume that every agent has a subjective prior—which may be different from the true distribution—and updates her belief in light of new information using Bayes’ law. By Berk’s theorem, any such agent whose prior is supported on the set of d -state models will forecast the observable in the long run *as if* she were using a pseudo-true d -state model. Focusing on pseudo-true models allows me to abstract from learning dynamics and focus on the asymptotic bias caused by misspecification.¹²

¹⁰Processes P and \tilde{P} are *observationally equivalent* if all their finite-dimensional marginal distributions are identical.

¹¹While Berk (1966) only covers the case of i.i.d. observations and parametric models, the result has been extended much more generally. Bunke and Milhaud (1998) and Kleijn and Van Der Vaart (2006) substantially extend Berk (1966) by providing conditions for the weak convergence of posterior distributions and considering infinite-dimensional models. Shalizi (2009)’s extension of Berk’s theorem covers the case of non-i.i.d. observations and hidden Markov models.

¹²One can alternatively consider agents who estimate the parameters of their d -state models using a quasi-maximum-likelihood estimator. Such agents also will asymptotically forecast *as if* they relied on pseudo-true d -state models. See, for instance, Theorem 2 of Douc and Moulines (2012).

The set of pseudo-true d -state models is independent of the agent’s preferences. Instead, it only depends on the number of states the agent can entertain and the true stochastic process. The independence of the agent’s pseudo-true models from her preferences is evident given the “as if” interpretation discussed above: Two agents who start with identical priors, observe the same sequence of observations, and update their beliefs using Bayes’ rule will end up with identical posteriors at any point in time—irrespective of their preferences. Berk’s theorem goes a step further by establishing that, in the long run, the posterior only depends on the support of the prior (not its other details) and the distribution of observations (not their realizations).

The independence of the agent’s pseudo-true models from her preferences has a significant consequence: The set of pseudo-true d -state models is generically disjoint from the set of d -state models that maximize the agent’s payoff. However, this disparity is a feature, not a bug, of a positive theory of bounded rationality. While finding the payoff-maximizing model requires knowledge of the true process, one arrives at the set of pseudo-true models simply by following Bayes’ rule—no knowledge of the true process is necessary. Following Bayes’ rule would have led the agent to the truth had her model been correctly specified, but it can lead her astray in the presence of model misspecification.

Thus, agents’ use of pseudo-true models should be seen as a positive statement—not a normative one. A pseudo-true d -state model is not what an agent should use for forecasting in order to maximize her payoff. It is what she will use to forecast in the long run if she starts with a prior over the set of d -state models and updates her belief using Bayes’ rule.

3 Pseudo-True Subjective Beliefs

In this section, I characterize the subjective belief of an agent who uses pseudo-true d -state models. As a preliminary step, I establish a useful invariance property for the class of pseudo-true models, which is of independent interest.

3.1 Linear Invariance

There are no constraints on the agent’s set of models other than the bound on the number of subjective state variables. Formally, matrices A , B , Q , and R of representation (1) are unrestricted, other than the minimal restrictions required for (1) to define a proper stationary ergodic stochastic process. This flexibility in the agent’s set of models enables her to capture any linear intratemporal relationship among observables by the appropriate choice of matrices A , B , Q , and R . It thus results in a crucial linear-invariance property for pseudo-true d -state models.

Theorem 2 (linear invariance). *Let $\tilde{y}_t = Ty_t$ denote a linear transformation of y_t , and let $\tilde{\mathbb{P}}$ denote the probability distribution over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$ induced by \mathbb{P} and T . Let \mathcal{P}_d^* denote the set of pseudo-true d -state stationary distributions when \mathbb{P} is the true process, and let $\tilde{\mathcal{P}}_d^*$ denote the corresponding set*

when $\tilde{\mathbb{P}}$ is the true process. If T is a full-rank matrix, then the set of probability distributions over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$ induced by \mathcal{P}_d^* and T coincides with $\tilde{\mathcal{P}}_d^*$.¹³

The theorem establishes that the agent's models and forecasts only depend on the amount of information available to her, not how it is presented. For instance, whether the agent observes the nominal interest rate and the inflation rate or the real interest rate and the inflation rate is immaterial for how she forms her expectations. Likewise, the agent's expectations are not affected by augmenting the vector of observables with linear combinations of variables already in her information set.

The dichotomy in the agent's cognitive abilities is arguably stark: The agent is capable of observing all the relevant variables and uncovering all linear *intratemporal* relationships among them, yet she is thoroughly constrained in the complexity of *intertemporal* relationships she can entertain. However, this dichotomy highlights the paper's premise that forecasting is challenging because it requires forecasters to recognize stochastic patterns that unfold over time. It also allows me to abstract away from the cognitive costs of acquiring information and the mistakes individuals make when information is presented differently. Furthermore, it makes the framework immediately portable across different applications, thanks to the linear-invariance result.¹⁴

The result also showcases the endogeneity of the agent's expectations. Since the parameters of the agent's model are determined endogenously by maximizing the fit to the true distribution, they covary with the true distribution. This feature of pseudo-true d -state models, which rational-expectations models share, makes the framework particularly suited to counterfactual analysis in macroeconomics, where policy changes can result in changes in the distribution of payoff-relevant variables.

The linear-invariance result allows me to focus on non-degenerate processes. Define the lag- l autocovariance matrix of the observable under the true process as follows:

$$\Gamma_l \equiv \mathbb{E}[y_t y_{t-l}']. \quad (2)$$

The true process is degenerate if Γ_0 is singular. Whenever Γ_0 is singular, there is some lower-dimensional vector \tilde{y}_t with $\mathbb{E}[\tilde{y}_t \tilde{y}_t']$ non-singular and some full-rank matrix T such that $y_t = T \tilde{y}_t$. By Theorem 2, the set of pseudo-true models when the observable is given by y_t can be found by first finding the set of pseudo-true models given \tilde{y}_t and then transforming that set by T .

Restricting attention to non-singular true processes allows me to restrict the agent to the set of models under which the subjective variance-covariance matrix of the observable is non-

¹³The distribution induced over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$ by a distribution P over $\{y_t\}_{t=-\infty}^{\infty}$ and a mapping $T : y_t \mapsto \tilde{y}_t$ is the pushforward distribution $P \circ T^{-1}$ defined as $P \circ T^{-1}(\{\tilde{y}_t\}_{t=-\infty}^{\infty} \in Y) \equiv P(\{y_t\}_{t=-\infty}^{\infty} : \{T y_t\}_{t=-\infty}^{\infty} \in Y)$ for any set $Y \subseteq \mathbb{R}^{n^Z}$. The set of distributions over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$ induced by a set of distributions \mathcal{P} over $\{y_t\}_{t=-\infty}^{\infty}$ and a mapping $T : y_t \mapsto \tilde{y}_t$ is given by $\mathcal{P} \circ T^{-1} \equiv \{P \circ T^{-1} : P \in \mathcal{P}\}$.

¹⁴Rabin (2013) calls for the use of portable extensions of existing models in behavioral economics and economic theory more generally. The current framework can be seen as a portable extension of the rational-expectations benchmark. By varying a single parameter, d , it spans the range between full rationality and a severe form of serial-correlation misperception where agents perceive serially-correlated variables as independent over time.

singular.¹⁵ In the remainder of the paper, I assume without loss that the variance-covariance matrix Γ_0 is non-singular, and the agent can only entertain subjective models with non-singular variance-covariance matrices.

3.2 The One-Dimensional Case

I start the analysis of the agent's pseudo-true models by considering the case where she can only entertain one-dimensional models. In this case, a complete characterization of the agent's forecasts is possible. The insights from the single-state case generalize to the d -state case, as I discuss later in this section.

The agent's pseudo-true one-state forecasts depend on the true process only through the autocorrelations of the vector of observables. Define the lag- l *autocorrelation matrix* of the observable under the true process as follows:

$$C_l \equiv \frac{1}{2} \Gamma_0^{-\frac{1}{2}} (\Gamma_l + \Gamma_l') \Gamma_0^{-\frac{1}{2}}. \quad (3)$$

The notion of autocorrelation matrices is a natural generalization of the notion of autocorrelation functions. When the observable y_t is a scalar, C_l reduces to the usual autocorrelation function at lag l . When the observable is an n -dimensional vector, on the other hand, C_l is an $n \times n$ real symmetric matrix with eigenvalues inside the unit circle.¹⁶

Autocorrelation matrices capture the extent of serial correlation in the vector of observables. Let $\rho(C_1)$ denote the spectral radius of matrix C_1 .¹⁷ When $\rho(C_l)$ is close to zero for all l , the process is close to being i.i.d., whereas when $\rho(C_l)$ is close to one, then the process is close to being unit root.

With the definition of autocorrelation matrices at hand, I can state the general characterization result for the $d = 1$ case:

Theorem 3. *Under any pseudo-true one-state model, the agent's s -period ahead forecast is given by*

$$E_t^{1*}[y_{t+s}] = a^s (1 - \eta) q p' \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}, \quad (4)$$

where a and η are scalars in the $[-1, 1]$ and $[0, 1]$ intervals, respectively, that maximize $\lambda_{\max}(\Omega(\tilde{a}, \tilde{\eta}))$, the largest eigenvalue of the $n \times n$ real symmetric matrix

$$\Omega(\tilde{a}, \tilde{\eta}) \equiv -\frac{\tilde{a}^2(1 - \tilde{\eta})^2}{1 - \tilde{a}^2 \tilde{\eta}^2} I + \frac{2(1 - \tilde{\eta})(1 - \tilde{a}^2 \tilde{\eta})}{1 - \tilde{a}^2 \tilde{\eta}^2} \sum_{\tau=1}^{\infty} \tilde{a}^\tau \tilde{\eta}^{\tau-1} C_\tau,$$

and $p = \Gamma_0^{-\frac{1}{2}} u$ and $q = \Gamma_0^{\frac{1}{2}} u$, where u is an eigenvector of $\Omega(a, \eta)$ with eigenvalue $\lambda_{\max}(\Omega(a, \eta))$, normalized so that $u'u = 1$.

¹⁵ Whenever the true variance-covariance matrix Γ_0 is non-singular, any subjective model with a singular variance-covariance matrix is dominated in terms of fit to the true process by every subjective model with a non-singular variance-covariance matrix. Therefore, no subjective model with a singular variance-covariance matrix can be a pseudo-true model.

¹⁶ See Lemma C.2 of the appendix for a proof.

¹⁷ The spectral radius $\rho(X)$ of matrix X denotes the maximum among the magnitudes of eigenvalues of X .

The endogenous variables a , η , p , and q have intuitive meanings. The scalar a captures the *persistence* of the subjective latent state. When $a = 0$, the subjective state is i.i.d., whereas when $a = 1$, it follows a unit-root process. The scalar η captures the *noise* in the agent's observations of the subjective state. When η is small, the agent believes recent observations to be highly informative of the value of the subjective state. As a result, her expectations respond more to recent observations and discount old observations more. The vector p determines the agent's *relative attention* to different components of the vector of observables. When p_i is larger than p_j , the agent puts more weight on $y_{i,t-\tau}$ relative to $y_{j,t-\tau}$ for all τ when forming her estimate of the subjective state. Finally, the vector q captures the *relative sensitivity* of the agent's forecasts of different observables to changes in her estimate of the subjective state. When q_i is larger than q_j , then a change in the estimated value of the state at time t leads the agent to change her forecast of $y_{i,t+s}$ by more than her forecast of $y_{j,t+s}$ for all s .

The theorem does not rule out the possibility that $|a| = 1$ and $\eta > 0$, in which case the corresponding state-space model would not be stationary ergodic. However, the following result establishes that any pseudo-true one-state model inherits the stationarity and ergodicity of the true process:

Theorem 4. *Let P^{1*} denote a pseudo-true one-state model given true distribution \mathbb{P} . If \mathbb{P} is purely non-deterministic and stationary ergodic, then so is P^{1*} .*

Theorem 3 significantly reduces the computational complexity of finding the set of pseudo-true models. The set of all d -state models is a non-compact manifold of dimension $2nd$ (Gevers and Wertz, 1984). Additionally, the KLD is a non-convex function of $\theta = (A, B, Q, R)$. Theorem 3 analytically concentrates out all but two of the parameters of the agent's models, thus reducing an optimization problem over a $2n$ -dimensional non-compact manifold to a problem over a two-dimensional compact rectangle.

Although much easier than the problem of KLD minimization over the space of one-state models, the problem of maximizing $\lambda_{\max}(\Omega(\tilde{a}, \tilde{\eta}))$ over $(\tilde{a}, \tilde{\eta})$ is still non-convex. Consequently, solving it requires the use of numerical global optimization methods. However, since the size of the problem is independent of n , it can be solved efficiently in any application, regardless of the dimension of the vector of observables.

The non-convexity of the problem also makes an analytical solution elusive without further assumptions on the true process. I proceed by imposing a natural assumption on the true process \mathbb{P} , which permits a closed-form characterization of the set of pseudo-true one-state models.

3.3 Exponential Ergodicity and Incomplete Information

The optimization problem in Theorem 3 has an intuitive closed-form solution given a class of true stochastic processes that arise naturally in many applications (including those studied in Sections 5–7). The appropriate class turns out to be the following:

Definition 1. A stationary ergodic process \mathbb{P} is *exponentially ergodic* if $\rho(C_l) \leq \rho(C_1)^l$ for all $l \geq 1$.¹⁸

Exponential ergodicity is stronger than ergodicity. While ergodicity requires the serial correlation at lag l to decay to zero as $l \rightarrow \infty$, exponential ergodicity requires the rate of decay to be faster than $\rho(C_1)$. Many standard processes are exponentially ergodic. For instance, the vector of observables follows an exponentially ergodic process if it is a linear combination of n independent AR(1) shocks.

The following result characterizes the agent's pseudo-true one-state forecasts when the true process is exponentially ergodic. It links the agent's forecasts to the eigenvalues and eigenvectors of the true autocorrelation matrix at lag one:

Theorem 5. *If the true process is exponentially ergodic, then the s -period ahead forecast of agents who use pseudo-true one-state models is given by*

$$E_t^{1*}[y_{t+s}] = a^s q p' y_t, \quad (5)$$

where a is an eigenvalue of C_1 largest in magnitude, u denotes the corresponding eigenvector normalized so that $u'u = 1$, and $p = \Gamma_0^{-\frac{1}{2}} u$ and $q = \Gamma_0^{\frac{1}{2}} u$.

Suppose there is a change in the value of the observable. The agent incorporates this information by first projecting the change in the observable on vector p to form an updated estimate $E_t^{1*}[z_t]$ of the subjective state. The agent thus dismisses as irrelevant any change in the vector of observables orthogonal to the relative attention vector p . She then forecasts the change in the s -period ahead value of the subjective state $E_t^{1*}[z_{t+s}]$ under the assumption that the state has persistence a . Finally, she multiplies her estimate of the subjective state by the relative sensitivity vector q to form her forecast of the observable in period $t + s$.

The following example illustrates the result in the context of a commonly-used specification for the true process:

Example 1. Suppose the true process \mathbb{P} has the following representation:

$$\begin{aligned} f_t &= F f_{t-1} + \epsilon_t, & \epsilon_t &\sim \mathcal{N}(0, \Sigma), \\ y_t &= H' f_t, \end{aligned}$$

where

$$F = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix},$$

¹⁸The term “exponential ergodicity” has been used to refer to a property of Markov chains, where the effect of initial condition on the current distribution of the state decays exponentially fast—see, for instance [Meyn and Tweedie \(1993\)](#). The definition used in this paper is mathematically distinct from the one in the context of Markov chains, but it captures the analogous idea that the serial correlation in the variables decays exponentially fast.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix},$$

$H \in \mathbb{R}^{n \times n}$ is an invertible square matrix, and $1 > |\alpha_1| > |\alpha_2| > \dots > |\alpha_n| > 0$. It is easy to verify that $\rho(C_l) = |\alpha_1|^l = \rho(C_1)^l$; that is, the true process is exponentially ergodic. Therefore, Theorem 5 can be used to characterize the pseudo-true one-state forecasts. The persistence, noise, relative attention, and relative sensitivity are, respectively, given by $a = \alpha_1$, $\eta = 0$, $p = (H'VH)^{\frac{1}{2}}H^{-1}V^{-1}e_1$, and $q = (H'VH)^{-\frac{1}{2}}H'Ve_1$, where $V \equiv (I - F^2)^{-1}\Sigma$ is the variance-covariance matrix of f_t and e_1 denotes the first coordinate vector.¹⁹

The agent's forecasts take a particularly simple form if H is the identity matrix, i.e., $y_{it} = f_{it}$ for $i = 1, \dots, n$. Then p and q are both multiples of the first coordinate vector e_1 , and the agent's forecasts simplify to

$$\begin{aligned} E_t^{1*}[y_{1,t+s}] &= \alpha_1^s y_{1t} = \mathbb{E}_t[y_{1,t+s}], \\ E_t^{1*}[y_{i,t+s}] &= 0, \quad \forall i \neq 1. \end{aligned}$$

The agent's forecast of the most persistent element of the vector of observables coincides with its rational-expectations counterpart, but she forecasts every other element of the observable as if it were i.i.d.

The example illustrates that the agent exhibits a form of *persistence bias*. She forecasts the most persistent component of the vector of observables as accurately as under rational expectations but misses the dynamics of other observables. The intuition for the result is easiest to see when the most persistent true state is close to being unit root. In that case, poorly tracking the most persistent state would lead to persistent mistakes in the agent's forecasts. The persistence of those mistakes would make them costly from the point of view of KLDL minimization. Therefore, any pseudo-true model tracks the state close to unit root as best possible, even if doing so results in errors in forecasting the other states. In Section 4.1, I generalize the insight of this example by formally establishing persistence bias in the context of general true processes.²⁰

A remarkable feature of the pseudo-true model in Example 1 is that the persistence parameter a does not depend on the volatilities of the underlying AR(1) processes. The agent uses the subjective latent state to track the most persistent component of y_t , even if the most persistent component has a vanishing variance. However, this result should not come as a surprise given the linear-invariance result: One can always equalize the volatilities of different components of y_t by an appropriate linear transformation of the observable without altering the persistence of

¹⁹See the proof of Lemma C.3 for a derivation.

²⁰Bidder and Dew-Becker (2016) and Dew-Becker and Nathanson (2019) propose an alternative reason agents might focus on tracking the most persistent components of a payoff-relevant variable. Bidder and Dew-Becker (2016) show that long-run risk is the worst case scenario for ambiguity-averse agents. Dew-Becker and Nathanson (2019) show that, as a result, ambiguity-averse agents will learn most about dynamics at the lowest frequencies.

the subjective latent state in the agent's pseudo-true model. Therefore, the persistence parameter cannot depend on the volatilities.

One can generalize Example 1 by relaxing the assumption that matrices F and Σ are diagonal and allowing for non-Gaussian innovations. The following theorem provides a set of sufficient conditions for the process to be exponentially ergodic:

Theorem 6. *Consider a true process \mathbb{P} that can be represented as*

$$\begin{aligned} f_t &= Ff_{t-1} + \epsilon_t \\ y_t &= H'f_t, \end{aligned} \tag{6}$$

where $f_t \in \mathbb{R}^m$, ϵ_t is a zero mean i.i.d. shock with a finite variance-covariance matrix, $F \in \mathbb{R}^{m \times m}$ is a convergent matrix, and $H \in \mathbb{R}^{m \times n}$. Suppose the variance-covariance of f_t is normalized to be the identity matrix. If H is a rank- m matrix and $\left\| \frac{F+F'}{2} \right\|_2 = \|F\|_2$, where $\|\cdot\|_2$ denotes the spectral norm, then the process is exponentially ergodic.

The assumption that the process has a representation of the form (6) is almost without loss of generality. By the Wold representation theorem, any mean zero, covariance stationary, and purely non-deterministic process can be approximated arbitrarily well by a process with a representation of the form (6). The assumption that the variance-covariance of f_t equals identity is also without loss of generality. It can always be arranged to hold by an appropriate normalization of f_t .²¹

The assumption on matrix F rules out a severe form of defectiveness by guaranteeing that the largest eigenvalue of the symmetric part of F coincides with the largest singular value of F . It is satisfied if F is diagonal or symmetric, for example. However, this assumption is much weaker than symmetry.

The most consequential assumption of the theorem is the requirement that H is a rank- m matrix. The assumption can be seen as a *complete information* (or spanning) assumption: If the agent observes an observable of the form (6) with a full-rank matrix H , then she has enough information to forecast the observable as well as in the full-information rational-expectations benchmark—even if she fails to do so due to the constraint on her set of models. The following proposition shows that this assumption, in general, cannot be dispensed with:

Proposition 1. *Suppose the observable is one-dimensional, and the true process \mathbb{P} can be represented as in (6) for some $f_t \in \mathbb{R}^m$, $\epsilon_t \sim N(0, \Sigma)$, diagonal divergent matrix $F \in \mathbb{R}^{m \times m}$, diagonal matrix $\Sigma \in \mathbb{R}^{m \times m}$, and matrix $H \in \mathbb{R}^{m \times n}$. If the representation in (6) is minimal and $m > 1$, then the s -period ahead forecast of agents who use pseudo-true one-state models is given by*

$$E_t^{1*}[y_{t+s}] = a^s(1 - \eta) \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}$$

for some $a \in (-1, 1)$ and $\eta \in (0, 1)$.

²¹See Lemma C.3 of the appendix and its proof for how this can be done.

The class of exponentially ergodic processes thus constitutes a subset of the class of stationary ergodic processes. However, the true processes that arise in the applications studied in this paper are all exponentially ergodic. In those applications, the forecasts of agents who are constrained to use pseudo-true one-state models take the simple form given by Theorem 5.

3.4 Pseudo-True d -Dimensional Models

I next investigate whether and how the results from the one-dimensional case generalize to the $d > 1$ case. The forecasts of agents who use d -state models take a form similar to equation (4): Given a d -state model $\theta = (A, B, Q, R)$, the agent's s -period ahead forecast is given by

$$E_t^\theta[y_{t+s}] = B'A^{s-1} \sum_{\tau=0}^{\infty} (A - KB')^\tau K y_{t-\tau}, \quad (7)$$

where $K \in \mathbb{R}^{d \times n}$ is the Kalman gain matrix, which depends on (A, B, Q, R) .²² Equation (7) is valid for any d -state model θ , not just the pseudo-true ones.

To characterize the forecasts under pseudo-true models, one needs to find the (A, B, Q, R) matrices that minimize the KLDR from the true process and the implied Kalman gain K . This is a hard problem that involves minimizing a non-convex function over the $2nd$ -dimensional non-compact manifold of d -state models Θ_d . While the problem can be simplified further, finding an analytical solution to the problem appears infeasible, even for $n = 1$.

I instead solve the easier problem of minimizing the KLDR over a subset $\bar{\Theta}_d$ of Θ_d . I say model $\theta = (A, B, Q, R)$ is *Markovian in observables* (m.i.o.) if $A - KB' = \mathbf{0}$, where K is the Kalman gain matrix corresponding to model θ and $\mathbf{0} \in \mathbb{R}^{d \times d}$ is the matrix of zeros. I let

$$\bar{\Theta}_d \equiv \{\theta = (A, B, Q, R) \in \Theta_d : A - KB' = \mathbf{0}\} \subset \Theta_d.$$

denote the set of m.i.o. d -state models. The time- t forecasts of an agent using a model $\theta \in \bar{\Theta}_d$ only depend on the realized value of the vector of observables at time t (and not its past realizations)—hence, the name Markovian in observables.

Model $\theta^* \in \bar{\Theta}_d$ is a pseudo-true m.i.o. d -state model if $\text{KLDR}(\theta^*) \leq \text{KLDR}(\theta)$ for all $\theta \in \bar{\Theta}_d$. I let $\bar{\Theta}_d^*$ denote the set of pseudo-true m.i.o. d -state models. The models in $\bar{\Theta}_d^*$ have several appealing theoretical properties. They satisfy a version of the linear-invariance result of Theorem 2, and they have similar Bayesian and quasi-maximum-likelihood learning foundations as (general) pseudo-true d -state models. Perhaps most importantly, the following corollary of Theorem 5 shows that constraining the agent to m.i.o. models is without loss under some conditions:

Corollary 1. *If the true process is exponentially ergodic, then any pseudo-true one-state model is m.i.o.*

The result shows that—at least in the one-dimensional case—the set of pseudo-true models is a subset of the set of m.i.o. models when the true process is exponentially ergodic. Whether

²²See equations (C.2) and (C.3) in the proof of Theorem 2 for the definition of K .

this result continues to hold for d -state models with $d > 1$ remains an open question. However, progress can be made by taking the restriction to m.i.o. models as an assumption and characterizing the set of pseudo-true m.i.o. d -state models and the corresponding forecasts:

Theorem 7. *Suppose the lag-one autocovariance matrix Γ_1 is symmetric. Then the s -period ahead forecast of agents who use pseudo-true m.i.o. d -state models is given by*

$$E_t^{d*}[y_{t+s}] = \sum_{i=1}^d a_i^s q_i p_i' y_t, \quad (8)$$

where a_1, \dots, a_d are the d eigenvalues of C_1 largest in magnitude (with the possibility that some of the a_i 's are equal), u_i denotes an eigenvector corresponding to a_i normalized such that $u_i' u_j = \delta_{ij}$ for all i and j , $p_i \equiv \Gamma_0^{-\frac{1}{2}} u_i$, $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$.

The result shows that the insights from the analysis of single-state models broadly carry over to d -state models. In particular, agents who are restricted to m.i.o. d -state models exhibit a form of persistence bias. They focus on perfectly forecasting the d most persistent elements of the vector of observables at the expense of the other elements.

The result suggests that state-space models can be estimated consistently by principal component analysis (PCA). This conclusion is reminiscent of a central result in the theory of dynamic factor models on the consistency of the principal components estimator for the common components.²³ However, Theorem 7 is different along several dimensions. First, it concerns state-space models, not dynamic factor models. Second, the estimator suggested by the theorem uses the principal components of the lag-one autocorrelation matrix, while the PCA estimator of dynamic factor models is constructed from the principal components of the variance-covariance matrix. Last but not least, Theorem 7 suggests that the PCA estimator is consistent (at least under the theorem's assumptions) even if the number of states is misspecified.²⁴ I am aware of no similar result on the consistency of the PCA estimator for dynamic factor models when the number of common factors is misspecified.

3.5 Second Moments

So far, the results of this section were concerned with the conditional first moments of pseudo-true d -state models. I end the section by presenting two results on the subjective second moments when the agent uses pseudo-true d -state models. The first result characterizes the agent's perceived variance-covariance of the vector of observables under pseudo-true one-state models.

Theorem 8. *Given any pseudo-true one-state model θ^* , the subjective variance-covariance of the vector of observables, $\text{Var}^{1*}(y_t)$, coincides with the true variance-covariance matrix, Γ_0 .*

²³See, for instance, [Stock and Watson \(2002\)](#).

²⁴An estimator for a misspecified model is consistent if the estimate converges to a pseudo-true model almost surely.

The agent does not misperceive the unconditional volatility of the vector of observables. Neither does she misperceive the unconditional means. Instead, it is the *conditional* expectations and volatilities that deviate from the corresponding values under rational expectations. The result is a direct consequence of the assumptions that the agent can entertain *any* stationary one-state model and fit her model to data by minimizing the KLDR. That the agent can entertain any one-state model allows her to always match the true volatility of the observables by an appropriate choice of matrices (A, B, Q, R) . The fact that the agent fits her model to data by minimizing the KLDR (or equivalently by Bayesian updating or maximum-likelihood estimation) means that it is optimal (from the point of view of maximizing fit) to match the volatility of the observable.

A version of this result also holds for d -state models:

Theorem 9. *Suppose the first autocovariance matrix Γ_1 is symmetric. Then given any pseudo-true m.i.o. d -state model θ^* , the subjective variance-covariance of the vector of observables, $\text{Var}^{d*}(y_t)$, coincides with the true variance-covariance matrix, Γ_0 .*

Agents who are constrained to use m.i.o. d -state models uncover the true variance-covariance matrix of the observable as long as the true process is sufficiently regular. This conclusion is a consequence of the fact that the set of m.i.o. models $\bar{\Theta}_d$ is invariant under linear transformations. For any Γ_0 , the set of m.i.o. models contains a subjective model θ such that $\text{Var}^\theta(y_t) = \Gamma_0$; KLDR minimization leads the agent to settle on such a subjective model.

4 Testable Predictions

I next use the characterization results of the previous section to develop several testable predictions of the framework for agents' forecasts and choices. When discussing agents' choices, I consider purely forward-looking actions that depend linearly on the agents' forecasts of the observable. In particular, I consider actions x_{jt} of the form

$$x_{jt} = E_t^j \left[\sum_{s=1}^{\infty} c'_{js} y_{t+s} \right], \quad (9)$$

where $j \in J$ indexes different actions (possibly taken by different agents), $y_t \in \mathbb{R}^n$ is as before the vector of observables, $E_t^j[\cdot]$ denotes the subjective forecasts of the agent taking action j , and $c_{jt} \in \mathbb{R}^n$ are constant vectors. I assume that $\sum_{s=1}^{\infty} \|c_{js}\|_2 < \infty$ for all j and that J is a finite set. I also continue to take the true process \mathbb{P} as a primitive of the economy. In Appendix B, I show that the predictions of the framework remain largely intact in a general equilibrium economy, in which \mathbb{P} itself is an endogenous outcome of agents' choices.

The reduced-form specification in (9) allows me to derive sharp results, which highlight the role of biased forecasts and are independent of the specifics of the agents' decision problems. These results hold up to first order for purely forward-looking decisions that depend non-linearly

on forecasts of the observable and arbitrarily well when decisions are sufficiently forward-looking (e.g., when the discount factor is close to one.) In the next three sections, I further develop the implications of the general framework in the context of three microfounded macro models.

4.1 Persistence Bias

A decomposition of y_t into its more persistent and less persistent components will prove fruitful for the discussions that follow.

Proposition 2. *Let a_i denote the i th largest eigenvalue of C_1 in magnitude, and let u_i denote the corresponding eigenvector, normalized such that $u_i' u_j = \delta_{ij}$ for all i and j . The vector of observables can be decomposed as follows:*

$$y_t = \sum_{i=1}^n y_t^{(i)} q_i, \quad (10)$$

where $y_t^{(i)} \equiv p_i' y_t$, $p_i \equiv \Gamma_0^{-\frac{1}{2}} u_i$, $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$, and scalars $y_t^{(i)}$ all have unit variance. If ρ_i denotes the lag-one autocorrelation of $y_t^{(i)}$, then $|\rho_1| \geq |\rho_2| \geq \dots \geq |\rho_n|$.

The proposition decomposes the vector of observables into the sum of n components: $\{y_t^{(i)} q_i\}_{i=1}^n$. The components are sorted in terms of their persistence, with $y_t^{(1)} q_1$ the most persistent and $y_t^{(n)} q_n$ the least persistent. This decomposition holds for arbitrary stationary stochastic process, without any reference to how agents forecast or what they do with their forecasts. However, the response of pseudo-true d -state forecasts to changes in the observable can be decomposed in a way that perfectly aligns with the decomposition in (10)

Agents who use pseudo-true d -state models treat the more and less persistent components of y_t in qualitatively different ways. Their expectations respond fully to the information contained in the d most persistent components of y_t , whereas they disregard the information content of the remaining components. The following corollary of Theorem 7 formalizes this idea:

Corollary 2 (persistence bias). *Suppose the lag-one autocovariance matrix Γ_1 is symmetric. Then the time t -forecasts of agents who use pseudo-true m.i.o. d -state models only depend on the d most persistent components of y_t .*

This persistence bias underpins the other testable predictions derived in this section. It leads to the unresponsiveness of expectations and actions, over-extrapolation of the most persistent components, under-extrapolation of the least persistent ones, and comovement among agents' different choices. In the remainder of this section, I discuss these predictions in turn.

4.2 Unresponsiveness on Impact

Ignoring innovations to the least persistent components of the observable makes the agents' forecasts and actions unresponsive to some new information. Consider a change in the current

value of the observable. The change can be decomposed into changes in components $\{y_t^{(i)} q_i\}_i$ of y_t . Agents do not change their forecasts in response to changes in the least persistent components of y_t . This unresponsiveness of expectations translates into the unresponsiveness of the agents' forward-looking actions, as established in the following corollary of Theorem 7:

Corollary 3. *Suppose the lag-one autocovariance matrix Γ_1 is symmetric, and consider an agent who uses pseudo-true m.i.o. d -state models and whose best responses are of the form (9). If $n > d$, the agent's time- t forecasts and actions do not respond to changes in the $n - d$ least persistent components of y_t .*

The agents' forecasts and actions are unresponsive to changes in $y_t^{(d+1)}, \dots, y_t^{(n)}$ only on impact. In general, different components of y_t do not evolve independently, so a change in the current value of $y_t^{(i)}$ could lead to changes in values of $y_{t+s}^{(j)}$ for some $j \neq i$ and $s > 0$. This could lead to a delayed response of the agents' forecasts and actions to changes in the least persistent components of the observable.

4.3 Underreaction and Overreaction

The framework proposed in this paper is neither a model of underreaction nor of overreaction. Instead, agents who use simple models forecast using a parsimonious model that provides an approximation to the true process and trades-off forecast errors at different horizons and for different variables. As such, pseudo-true simple models do not lead to mistakes that always go in the "same direction." In fact, agents who use pseudo-true one-state models always under-extrapolate some variables and over-extrapolate others:

Proposition 3. *Let $y_t^{(1)}$ denote the most persistent component of y_t and $y_t^{(n)}$ denote its least persistent component. Suppose the true process for y_t is exponentially ergodic. Consider an agent who uses a pseudo-true one-state model.*

- (a) *The agent overestimates the magnitude of $y_t^{(1)}$'s autocorrelation at all lags.*
- (b) *If $n > 1$, the agent underestimates the magnitude of $y_t^{(n)}$'s autocorrelation at all lags.*

The following example illustrates this result:

Example 2. Suppose the vector of observables is given by $y_t = (y_{1t}, y_{2t})' \in \mathbb{R}^2$, and each element of y_t follows an independent ARMA(1, 1) process

$$\begin{aligned} y_{1t} &= \phi_1 y_{1,t-1} + \epsilon_{1t} + \vartheta_1 \epsilon_{1,t-1}, \\ y_{2t} &= \phi_2 y_{2,t-1} + \epsilon_{2t} + \vartheta_2 \epsilon_{2,t-1}, \end{aligned}$$

where $\phi_1, \phi_2, \vartheta_1, \vartheta_2 \in (0, 1)$ are constants, and ϵ_{1t} and ϵ_{2t} are i.i.d. mean-zero random variables with finite variances. Additionally, assume that $\phi_1 > \phi_2$ and $\frac{(\phi_1 + \vartheta_1)(1 + \phi_1 \vartheta_1)}{1 + 2\phi_1 \vartheta_1 + \vartheta_1^2} > \frac{(\phi_2 + \vartheta_2)(1 + \phi_2 \vartheta_2)}{1 + 2\phi_2 \vartheta_2 + \vartheta_2^2}$. This assumption ensures that y_{1t} has a higher autocorrelation than y_{2t} at all lags.

The lag- l autocorrelation matrix is given by

$$C_l = \begin{pmatrix} \frac{(\phi_1 + \vartheta_1)(1 + \phi_1 \vartheta_1)}{1 + 2\phi_1 \vartheta_1 + \vartheta_1^2} \phi_1^{l-1} & 0 \\ 0 & \frac{(\phi_2 + \vartheta_2)(1 + \phi_2 \vartheta_2)}{1 + 2\phi_2 \vartheta_2 + \vartheta_2^2} \phi_2^{l-1} \end{pmatrix}.$$

The i th largest eigenvalue of C_1 in magnitude is $\frac{(\phi_i + \vartheta_i)(1 + \phi_i \vartheta_i)}{1 + 2\phi_i \vartheta_i + \vartheta_i^2}$, and the corresponding eigenvector is $u_i = e_i$, where e_i denotes the i th standard coordinate vector. Since the two elements of y_t are independent, the variance-covariance matrix Γ_0 is diagonal. Therefore, $y_t^{(1)} q_1 = y_{1t} e_1$ and $y_t^{(2)} q_2 = y_{2t} e_2$, i.e., the most persistent component of y_t is its first element and the least persistent component of y_t is its second element. It is also easy to verify that

$$\rho(C_l) = \frac{(\phi_1 + \vartheta_1)(1 + \phi_1 \vartheta_1)}{1 + 2\phi_1 \vartheta_1 + \vartheta_1^2} \phi_1^{l-1} \geq \left(\frac{(\phi_1 + \vartheta_1)(1 + \phi_1 \vartheta_1)}{1 + 2\phi_1 \vartheta_1 + \vartheta_1^2} \right)^l = \rho(C_1)^l,$$

with the inequality strict for $l > 1$. That is, the true process is exponentially ergodic.

The pseudo-true one-state model is described by Theorem 5. Under any such model, y_{1t} follows an AR(1) process with persistence parameter $a = \frac{(\phi_1 + \vartheta_1)(1 + \phi_1 \vartheta_1)}{1 + 2\phi_1 \vartheta_1 + \vartheta_1^2}$ and y_{2t} is i.i.d. over time. The pseudo-true lag- l autocorrelation of y_{1t} is equal to a^l , while the pseudo-true lag- l autocorrelation of y_{2t} is zero for any $l \geq 1$. On the other hand, the true lag- l autocorrelation of y_{it} is given by $\frac{(\phi_i + \vartheta_i)(1 + \phi_i \vartheta_i)}{1 + 2\phi_i \vartheta_i + \vartheta_i^2} \phi_i^{l-1}$ for $i = 1, 2$. Therefore, an agent using a pseudo-true one-state model overestimates the autocorrelation of y_{1t} at all lags (strictly so for lags $l > 1$) while strictly underestimating the autocorrelation of y_{2t} at all lags.

The overreaction/underreaction result generalizes to the $d > 1$ case by imposing a slightly stronger assumption on the true process and focusing on models that are Markovian in the observable:

Proposition 4. *Let $y_t^{(1)}$ denote the most persistent component of y_t and $y_t^{(n)}$ denote its least persistent component. Suppose the true process for y_t is exponentially ergodic and that the lag-one autocovariance matrix Γ_1 is symmetric. Consider an agent who uses a pseudo-true m.i.o. d -state model with $d < n$.*

- (a) *The agent overestimates the magnitude of $y_t^{(1)}$'s autocorrelation at all lags.*
- (b) *The agent underestimates the magnitude of $y_t^{(n)}$'s autocorrelation at all lags.*

Relative to the rational-expectations benchmark, agents who use simple models overreact to changes in the most persistent components of the observable and underreact to changes in the least persistent ones. For observables with intermediate persistence, the pattern could be underreaction or overreaction depending on the variable and the horizon being considered. These predictions set this paper's framework apart from models that generate only underreaction or only overreaction.

4.4 Comovement

Forward-looking actions of agents who are constrained to use simple models comove more than under rational expectations. This result might seem like a mechanical consequence of constraining agents to use a small number of state variables. If $d = 1$, for example, each decision depends on a single component of y_t , so it might appear obvious that agents' different decisions should perfectly comove. However, this conclusion would be false if different decisions depended on different components of y_t .

The comovement between different choices relies crucially on the assumption that agents use pseudo-true models. Since the set of pseudo-true d -state models is independent of the specifics of the decision being considered, not only does each decision respond only to d components of y_t , but also different decisions respond to the exact same d components. In fact, even agents with different d 's respond to overlapping subsets of a small set of common components. Therefore, different actions made by different agents with different d 's all vary with movements in a small set of common components. The next result formalizes this intuition:

Proposition 5. *Consider actions x_{jt} of the form (9) chosen by possibly different agents who use pseudo-true m.i.o. d -state models with possibly different d 's. Let D denote the largest value of d across agents and $x_t \equiv (x_{jt})_{j \in J} \in \mathbb{R}^{|J|}$ denote the vector containing the agents' time- t actions. Suppose the lag-one autocovariance matrix Γ_1 is symmetric and the eigenvalues of the lag-one autocorrelation matrix C_1 are all distinct. If $D \leq |J|$, then there exists a D -dimensional subspace \mathcal{V} of $\mathbb{R}^{|J|}$ such that $x_t \in \mathcal{V}$ for all t .*

The proposition establishes that the agents' actions exhibit a strong form of comovement. Not only do different actions of each agent move with one another, but also actions of different agents comove. To see the intuition for this result, consider agents j and j' who use models of dimensions $d(j)$ and $d(j') > d(j)$, respectively. The agents disagree on how many components of y_t are needed to forecast the observable, but they agree on what the $d(j)$ most important components are. The $d(j)$ components used by agent j are a subset of the $d(j')$ components used by agent j' . This strong form of comovement is a unique prediction of the framework.

The assumption that the eigenvalues of the lag-one autocorrelation matrix are all distinct is satisfied for generic true processes—even within the class of processes for which the lag-one autocovariance matrix is symmetric. This assumption implies that all pseudo-true models of a given dimension are observationally equivalent. It rules out knife-edge cases where several substantially different models of a given dimension have the exact same, minimal Kullback–Leibler divergence rate from the true process.

When agents are all constrained to use one-state models, their different actions all perfectly align:

Proposition 6. *Consider actions x_{jt} of the form (9) chosen by possibly different agents who all use pseudo-true one-state models. Let $x_t \equiv (x_{jt})_{j \in J} \in \mathbb{R}^{|J|}$ denote the vector containing the agents'*

time- t actions. Suppose the true process is exponentially ergodic, and the top eigenvalue of the lag-one autocorrelation matrix C_1 is simple. Then there exists a 1-dimensional subspace \mathcal{V} of $\mathbb{R}^{|J|}$ such that $x_t \in \mathcal{V}$ for all t .

The time- t actions of agents who use pseudo-true one-state models only depend on the current realization of $y_t^{(1)}$, the most persistent component of y_t . Consequently, an econometrician who analyzes those actions will conclude that the actions are driven by a single “main shock.” This conclusion is independent of the specifics of preferences, technology, or market structure. It holds both in partial equilibrium and in general equilibrium, as is suggested by the analysis in Appendix B.

Angeletos, Collard, and Dellas (2020) find that a “main business cycle shock” explains the bulk of movements in macroeconomic aggregates at business cycle frequencies. Based on my analysis, this finding should not come as a surprise: There is always a main shock—as long as decisions are sufficiently forward looking and agents use simple models. But in general, the main shock is an endogenous index whose composition depends on the primitives of the economy such as preferences, technology, market structure, the stochastic properties of the shocks that hit the economy, and the parameters of policy rules.

The subspace-based notion of comovement is special, but it is the appropriate notion in this context. An alternative, commonly used measure of comovement between two variables is their Pearson correlation coefficient. If comovement is measured by the correlation coefficients between different actions, constraining agents to d -state models might lower comovement relative to the rational expectations benchmark. However, the following corollary of Proposition 6 shows that in the $d = 1$ case comovement increases, even when measured by Pearson’s correlation coefficient:

Corollary 4. *Consider actions j and k , both of the form (9). Suppose the true process is exponentially ergodic, and the top eigenvalue of the lag-one autocorrelation matrix C_1 is simple. Then*

$$1 = \left| \text{Corr}\left(x_{jt}^{1*}, x_{kt}^{1*}\right) \right| \geq \left| \text{Corr}\left(x_{jt}^{RE}, x_{kt}^{RE}\right) \right|,$$

with the inequality generically strict.

5 Application to the New-Keynesian Model

As the first application of the general framework, I study the standard three-equation new-Keynesian model.²⁵

²⁵Technically speaking, the economy will be a two-equation new-Keynesian economy, described by the dynamic IS curve and the Phillips curve. Instead of assuming a functional form for the way the interest rate is set (e.g., a Taylor rule), I model the nominal interest rate as a random variable with arbitrary dependence on current realizations and lags of output gap and inflation rate.

5.1 Primitives

The primitives of the economy are standard. Time is discrete, preferences are time separable, and discounting is exponential. There is a measure of households with separable preferences over the final good and leisure. In each period, households decide how much to consume and how much to save in a nominal bond, which is in zero net supply. Households also make labor-supply decisions taking the wage as given. The consumption good is a CES aggregate of a continuum of intermediate goods. Intermediate goods are produced by monopolistically competitive firms using a technology linear in labor. Intermediate-good producers are subject to a Calvo-style pricing friction. Markets for labor, the final good, and the nominal bond are competitive.

The economy is subject to technology shocks that move the natural rate of interest and cost-push shocks that affect the intermediate-good producers' desired markups. The nominal interest rate is set by a central bank. The exact rule followed by the central bank is irrelevant for my analysis. Rather, equilibrium outcomes will depend only on the statistical properties of the interest rate process (such as its serial correlation and its correlation with other aggregate observables).²⁶

5.2 Log-linear Temporary Equilibrium

It is well known since [Preston \(2005\)](#) that recursive equilibrium equations that relate aggregate variables (e.g., the aggregate Euler equation) may not be valid outside of rational expectations. Instead, one needs to separately characterize each agent's optimal behavior using only relationships that are respected by the agent's expectations.

My analysis of the new-Keynesian model thus proceeds in two steps. The first step is to characterize the *temporary equilibrium* relationships, which impose individual optimality and market clearing conditions but not rational expectations.²⁷ The second step is to supplement the temporary equilibrium with the model of expectations formation and characterize the resulting (full) equilibrium.

The first step of the analysis is standard. I therefore omit the details of the derivation and use the log-linearized temporary equilibrium relationships as my starting point.²⁸ These temporary-equilibrium conditions are given by

$$\hat{x}_t = -\sigma (\hat{l}_t - r_t^n) + E_t^h \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{l}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right], \quad (11)$$

$$\hat{\pi}_t = \kappa \hat{x}_t + \mu_t + E_t^f \left[\sum_{s=1}^{\infty} (\beta \delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right], \quad (12)$$

²⁶The Taylor principle is not necessary for equilibrium uniqueness in my setup. The sunspot equilibria in the new-Keynesian model require agents to observe payoff-irrelevant common sunspots. I restrict the set of variables that can appear in the vector of observables to be payoff relevant, thus implicitly ruling out sunspot equilibria.

²⁷The notion of temporary equilibrium has been extensively developed in the context of Arrow–Debreu economies by [Grandmont \(1977\)](#). See [Woodford \(2013\)](#) for a discussion of temporary equilibria in the context of modern monetary models and [Farhi and Werning \(2019\)](#) for an application in the context of a heterogeneous-agent new-Keynesian economy.

²⁸Details of this derivation can be found, among other places, in [Angeletos and Lian \(2018\)](#) and [Gáti \(2020\)](#).

where \hat{x}_t , \hat{i}_t , and $\hat{\pi}_t$ denote the log-deviations of output gap, (gross) nominal interest rate, and inflation rate, respectively, from their steady state values. β is the discount factor, σ is the elasticity of intertemporal substitution (EIS), δ is the Calvo parameter, κ is a composite parameter that determines the steepness of the Phillips curve, r_t^n denotes the technology shock that moves the natural rate of interest, and μ_t is the cost-push shock. E_t^h and E_t^f denote subjective expectations of households and firms, respectively.

I assume that the vector $(\hat{i}_t, r_t^n, \mu_t)'$ of nominal interest rate, technology shock, and cost-push shock follows a mean-zero stationary and exponentially ergodic process.²⁹ This assumption allows me to use Theorem 5 to characterize the set of pseudo-true one-state models. When taking the model to data, I verify that $(\hat{i}_t, r_t^n, \mu_t)'$ indeed follows a stationary and exponentially ergodic process.

5.3 Subjective Expectations and Equilibrium

For simplicity, I assume that households and firms face identical constraints on the models they can entertain, ending up with identical subjective expectations. Every agent knows the steady-state values of every variable. The agents' time- t information set is given by the history $\{y_\tau\}_{\tau \leq t}$ of vector $y_\tau \equiv (\hat{x}_\tau, \hat{\pi}_\tau, \hat{i}_\tau, r_\tau^n, \mu_\tau)'$, consisting of time- τ log-deviations of output, inflation, and interest rate from their steady-state values, as well as realizations of every shock. Instead of imposing rational expectations, I assume agents are constrained to use one-dimensional state-space models of form (1) to forecast y .

The equilibrium definition is straightforward. An equilibrium consists of a stochastic process \mathbb{P} for $\{y_t\}_t$ and a model θ^* for agents such that (i) \mathbb{P} is derived from market-clearing conditions and optimal behavior by households and firms given subjective model θ^* , and (ii) θ^* is a pseudo-true one-state model given the stochastic process \mathbb{P} . Following earlier work (Molavi, 2019), I refer to this equilibrium notion as *constrained rational expectations equilibrium*.

Finding an equilibrium involves solving a fixed-point equation. I can do this in the context of the new-Keynesian model analytically via a guess-and-verify method. I focus on linear equilibria, in which \hat{x}_t and $\hat{\pi}_t$ are linear functions of \hat{i}_t , r_t^n , and μ_t .³⁰ In such an equilibrium, the y_t vector contains two redundant elements (which are linear combinations of other elements of y_t). Therefore, agents' forecasts of y can be obtained by first finding their forecasts of some three-dimensional vector f that spans the subspace spanned by y and then using the linear-invariance result to find their forecasts of y .

²⁹The new-Keynesian literature often assumes that nominal interest rate follows a Taylor rule, which sets the rate as a linear function of output gap and inflation rate plus a monetary policy shock. As long as shocks follow a stationary and exponentially ergodic process, the standard specification leads to a process for $(\hat{i}_t, r_t^n, \mu_t)'$ that is stationary and exponentially ergodic—both in the rational expectations equilibrium and in the equilibrium in which agents are constrained to use simple state-space models. My reduced-form specification of the interest rate process thus nests the standard Taylor-rule specification. But the reduced-form specification has the advantage of allowing the model to be calibrated without taking a stand on which changes in the interest rate are systematic and which are due to the so-called pure monetary policy shocks. It also enables me to study the effects of forward guidance in a theoretically coherent way. These advantages come at the expense of precluding counterfactual analyses with respect to the parameters of the Taylor rule.

³⁰The existence and generic uniqueness of a linear equilibrium follows from the guess-and-verify argument. My method for finding an equilibrium is silent on whether there are other, non-linear equilibria.

I take $f_t \equiv (\hat{x}_t, \hat{\pi}_t, \hat{i}_t)'$ as my basis for the subspace spanned by y_t . This choice of f_t has two advantages over the more natural choice of the vector of shocks. First, it considerably simplifies the algebra involved in finding the equilibrium. Second, it makes the estimation of the model more straightforward. By Theorem 5, the agents' pseudo-true model of any vector f_t depends on the autocovariance matrices of f_t at lags zero and one. When f_t consists of output gap, inflation, and interest rate, those autocovariance matrices have readily available empirical counterparts.

The following proposition summarizes the equilibrium characterization:

Proposition 7. *Suppose the shocks in the new-Keynesian model are stationary and exponentially ergodic, agents are constrained to use one-state models, and their time- t information set consists of the history of vector $y_\tau \equiv (\hat{x}_\tau, \hat{\pi}_\tau, \hat{i}_\tau, r_\tau^n, \mu_\tau)'$ for $\tau \leq t$. In the linear constrained rational expectations equilibrium,*

$$\hat{x}_t = \frac{1}{1 - p_x \gamma_x - p_\pi (\gamma_\pi + \kappa \gamma_x)} \left[\gamma_x (p_i \hat{i}_t + p_\pi \mu_t) - \sigma (1 - \gamma_\pi p_\pi) (\hat{i}_t - r_t^n) \right], \quad (13)$$

$$\hat{\pi}_t = \frac{1}{1 - p_x \gamma_x - p_\pi (\gamma_\pi + \kappa \gamma_x)} \left[(\gamma_\pi + \kappa \gamma_x) p_i \hat{i}_t + (1 - \gamma_x p_x) \mu_t - \sigma (\kappa + \gamma_\pi p_x) (\hat{i}_t - r_t^n) \right], \quad (14)$$

where

$$\gamma_x \equiv a(q_x - \sigma q_\pi), \quad (15)$$

$$\gamma_\pi \equiv a\beta q_\pi, \quad (16)$$

Γ_0 is the variance-covariance matrix of $(\hat{x}_t, \hat{\pi}_t, \hat{i}_t)$, C_1 is the corresponding lag-one autocorrelation matrix, a is the eigenvalue of C_1 largest in magnitude, u is the corresponding eigenvector normalized so that $u'u = 1$, $p \equiv (p_x, p_\pi, p_i)' \equiv \Gamma_0^{-\frac{1}{2}} u$, and $q \equiv (q_x, q_\pi, q_i)' \equiv \Gamma_0^{\frac{1}{2}} u$.

The proposition provides an explicit characterization of the equilibrium given autocovariance matrices of vector $f_t = (\hat{x}_t, \hat{\pi}_t, \hat{i}_t)'$. Although f_t contains output gap and inflation rate, which are endogenous objects, the characterization is still useful. One can directly measure the autocovariance matrices of f_t in the data and use the measured values together with values for β , σ , δ , and κ to find the response of the economy to interest rate changes as well as technology and cost-push shocks. Furthermore, in equilibrium, there is a one-to-one mapping between autocovariance matrices of f_t and autocovariance matrices of the shocks. Therefore, setting the autocovariance matrices of f_t to their empirical counterparts is equivalent to choosing the shock process to target the empirical autocovariance matrices of f_t .

Vectors p and q can be seen as measures of the anchoring of expectations. The agents' nowcast of the subjective latent state z_t can be seen as their view of the "state of the economy." When p_ζ is small for some $\zeta \in \{x, \pi, i\}$, the agents' view of the state of the economy does not move by much in response to innovations in ζ_t . Whereas when q_ζ is small for some $\zeta \in \{x, \pi, i\}$, changes in the agents' view of the state of the economy do not alter their forecasts of ζ by much. The product $q_\zeta p_\zeta$ thus captures the sensitivity of forecasts of ζ' to innovations in ζ . When q_ζ is small,

the agents' expectations of ζ are well anchored; they do not respond to innovations in any of the observables.

The framework can be used to think about monetary policy when agents use simplified models of the economy. The standard new-Keynesian model has the so-called “divine coincidence” property under rational expectations: Without cost-push shocks, the monetary authority faces no trade-off between its dual goals of zero output gap and stable prices. The following result shows that a similar conclusion holds when agents use pseudo-true one-state models:

Proposition 8 (divine coincidence). *Suppose agents in the new-Keynesian model are constrained to use one-state models, and their time- t information set consists of the history of vector $y_\tau \equiv (\hat{x}_\tau, \hat{\pi}_\tau, \hat{i}_\tau, r_\tau^n, \mu_\tau)'$ for $\tau \leq t$. If the cost-push shock μ_t is identically zero, the nominal interest rate is always set to the natural rate, i.e., $\hat{i}_t = r_t^n$, and the natural rate follows a stationary and exponentially ergodic process, then in the linear constrained rational expectations equilibrium, the output gap and inflation rate are identically zero.*

Equalizing the nominal rate and the natural rate achieves zero output gap and inflation through two channels. The first channel is the direct effect on the current interest rate faced by households and firms. The second channel works through the anchoring of inflation expectations. By systematically setting the nominal rate to the natural rate, the monetary authority brings the economy closer to a one-state economy, in which the only shocks affecting the economy are shocks to the natural rate. This brings the agents' pseudo-true one-state expectations closer to rational expectations, thus making the constraint on the complexity of the agents' models non-binding. As a result, the economy inherits the divine coincidence property of the rational expectations version of the model. In equilibrium, $q_x = q_\pi = 0$; that is, the agents' expectations of output gap and inflation are perfectly anchored.

Limits to the complexity of agents' models does not change the prescription for conventional monetary policy (at least when cost-push shocks are absent). But it is a completely different story when it comes to forward guidance. I proceed by quantifying the effect of forward guidance in an economy where agents use simple models fit to their past observations.

5.4 Forward Guidance

I consider an economy that has been operating without forward guidance for a long time and study how implementing forward guidance then affects output and inflation. This is a good description of where the U.S. economy was in 2009, in the aftermath of the Global Financial Crisis. Consistent with this story, I end my sample in the fourth quarter of 2008 when taking the model to data.

I assume that agents continue to forecast using a one-state model that is pseudo true in an equilibrium without forward guidance even as they see forward guidance. This assumption captures the following scenario: Agents have lived in a new-Keynesian economy without forward

guidance for a long time and have had ample opportunities to learn the equilibrium relationships. However, since agents can only entertain one-state models, instead of learning the true model, they have settled on a pseudo-true one-state model. Agents are then confronted with forward guidance for the first time. The key assumption is that agents do not immediately abandon their model; rather, they continue to rely on the model they had before the switch to the forward-guidance regime, even though the model may not be pseudo-true under the new regime.

The fact that agents have a fully specified model for the stochastic process of y allows me to study the effects of forward guidance in an internally consistent way. I model forward guidance as a credible announcement in period t by the central bank that the nominal rate will follow path $\{\hat{i}_{t+1}, \hat{i}_{t+2}, \dots, \hat{i}_{t+T}\}$ going forward. The announcement augments the agents' time- t information set to include $\{\hat{i}_{t+1}, \hat{i}_{t+2}, \dots, \hat{i}_{t+T}\}$ (in addition to $\{y_\tau\}_{\tau \leq t}$). Therefore, the agents' time- t forecasts under forward guidance are the conditional expectations $E_{t,FG(T)}^{1*}[\cdot] \equiv E_t^{1*}[\cdot | \{y_\tau\}_{\tau \leq t}, \hat{i}_{t+1}, \hat{i}_{t+2}, \dots, \hat{i}_{t+T}]$. But the agents' forecasts are Markovian in observables by Theorem 5 and the assumption that the true process is exponentially ergodic. Therefore, $E_{t,FG(T)}^{1*}[\cdot] = E_t^{1*}[\cdot | y_t, \hat{i}_{t+1}, \hat{i}_{t+2}, \dots, \hat{i}_{t+T}]$.

Since agents use linear-Gaussian state-space models, their forecasts are linear functions of the variables in their information set. In particular, for any observable $\zeta \in \{\hat{x}, \hat{\pi}, \hat{i}, r^n, \mu\}$

$$E_{t,FG(T)}^{1*}[\zeta_{t+s}] = E^{1*}[\zeta_{t+s} | f_t, \hat{i}_{t+1}, \hat{i}_{t+2}, \dots, \hat{i}_{t+T}] = \Sigma_{\zeta_s \omega_T} \Sigma_{\omega_T \omega_T}^{-1} \omega_T,$$

where $\omega_T \equiv (\zeta_t, \hat{i}_{t+1}, \dots, \hat{i}_{t+T})'$, $\Sigma_{\zeta_s \omega_T} \equiv E^{1*}[\zeta_{t+s} \omega_T']$, and $\Sigma_{\omega_T \omega_T} \equiv E^{1*}[\omega_T \omega_T']$. Note that the covariance matrices that show up in the agents' forecasts of ζ are subjective covariance matrices which depend on the agents' subjective model. But the subjective model is just the pseudo-true one-state model, which is fully characterized by Proposition 7.

The response of the economy to forward guidance takes a relatively simple form. Substituting for the agents' forecasts in (11) and (12) and simplifying the resulting expression, I obtain

$$\begin{aligned} \hat{x}_t &= v_{xi}^{(T)} \hat{i}_t + v_{xn}^{(T)} r_t^n + v_{x\mu}^{(T)} \mu_t + \sum_{s=1}^T v_{xi_s}^{(T)} \hat{i}_{t+s}, \\ \hat{\pi}_t &= v_{\pi i}^{(T)} \hat{i}_t + v_{\pi n}^{(T)} r_t^n + v_{\pi \mu}^{(T)} \mu_t + \sum_{s=1}^T v_{\pi i_s}^{(T)} \hat{i}_{t+s}, \end{aligned}$$

where v 's are constants that depend on the parameters (a, p, q) of the agents' pseudo-true model and constants β, σ, δ , and κ . The expressions for v 's can be found in Online Appendix D.1.

The v 's have intuitive interpretations: v_{xi} and $v_{\pi i}$ are the current interest rate elasticities of output and inflation, respectively, whereas v_{xi_s} and $v_{\pi i_s}$ are the elasticities of output and inflation with respect to the s -period-ahead interest rate. Note that these elasticities change with the duration T of the monetary authority's guidance. That is, committing to a zero interest rate in period $t + s$ is *not* the same as not making any announcement about period $t + s$'s interest rate. The (T) superscript in the above expressions emphasizes this point. The expressions for v 's are rather cumbersome and hard to interpret, so I instead calibrate the model and numerically study the effects of forward guidance.

5.5 Calibration and Estimation

The model has few parameters. I calibrate the model at a quarterly frequency. Following Galí (2015), I set $\beta = 0.99$, $\sigma = 1$, $\delta = 3/4$, and $\kappa = 0.172$. I choose the first two autocovariance matrices of vector $(\hat{i}_t, r_t^n, \mu_t)'$ of nominal rate, technology shock, and cost-push shock to match the first two autocovariance matrices of $f_t = (\hat{x}_t, \hat{\pi}_t, \hat{i}_t)'$ (in a constrained rational expectations equilibrium where agents use one-state models). Since there is a one-to-one mapping between the two sets of autocovariance matrices, I can perfectly match the autocovariance matrices of f_t .

I estimate the empirical autocovariance matrices of f_t using the post-war, pre-Global-Financial-Crisis U.S. data. For \hat{x}_t , I use the percentage difference between real GDP and Potential Output in period t ; for $\hat{\pi}_t$, I use the percentage change in GDP Deflator; and for \hat{i}_t , I use the Effective Fed Funds Rate. The resulting time series are stationary, so I do not filter them. The sample period is from the first quarter of 1955 to the fourth quarter of 2008.

The estimated (lag-one) autocorrelations of interest rate, technology shock, and cost-push shock are given, respectively, by $\rho_i = 0.954$, $\rho_{r^n} = 0.955$, and $\rho_\mu = 0.925$, whereas the corresponding standard-deviations are given by $\sigma_i = 3.30$, $\sigma_{r^n} = 5.67$, and $\sigma_\mu = 0.315$. However, the estimated shocks are not independent AR(1) processes. See Online Appendix D.2 for the full estimated autocovariance matrices at lags zero and one, where I also verify that the estimated process is exponentially ergodic.

There are no free parameters for agents' expectations. Agents' models, beliefs, and forecasts are all pinned down by structural parameters β , σ , δ , and κ and the stochastic process of the shocks. In equilibrium, the agents' pseudo-true one-state model is described by

$$\begin{aligned} a^* &= 0.985, \\ p_x^* &= 0.022, \\ p_\pi^* &= -0.42, \\ p_i^* &= -0.014, \\ q_x^* &= 0.53, \\ q_\pi^* &= -2.3, \\ q_i^* &= -2.5, \end{aligned}$$

where a^* denotes the perceived persistence, p^* is the relative attention vector, and q^* is the relative sensitivity vector.

Agents perceive the subjective state, i.e., “the state of the economy,” as highly persistent but not unit root. Their estimate of the state of the economy is highly sensitive to changes in inflation, but it does not respond much to output or interest rate. High output makes agents optimistic about the state of the economy, while high inflation and high interest rate make them pessimistic. Finally, agents' forecasts of inflation and interest rate move considerably with their estimate of the state of the economy, but not so much for output.

The inflation expectations of an agent using the pseudo-true one-state model are not well anchored. A one percent increase in the current inflation rate would result in a $a^* q_\pi^* p_\pi \approx 0.95\%$ increase in the agent's one-period ahead forecast of inflation rate. The non-anchoring of expectations reflects the large persistence of inflation in the time-series data used to estimate the model.

5.6 Results

Figure 1 plots the responses of output and inflation to a 100 basis point cut in the current nominal rate combined with an announcement by the central bank that the nominal rate will be kept at -1% for T quarters. The figure plots the response at the time of announcement as the duration of guidance T is varied. The response of output to a 100 basis point rate cut accompanied by a promise to keep the rate low for another quarter is almost 50% higher than the response to a rate cut without any guidance. But the central bank quickly runs out of ammunition. Promising to keep the rate low for two quarters (instead of one) increases the response of output only by about 9%. Promising to keep the rate low for 20 quarters is only 50% more stimulative than promising to keep it low for one quarter. Likewise, the response of inflation to forward guidance does not change by much as T increases.

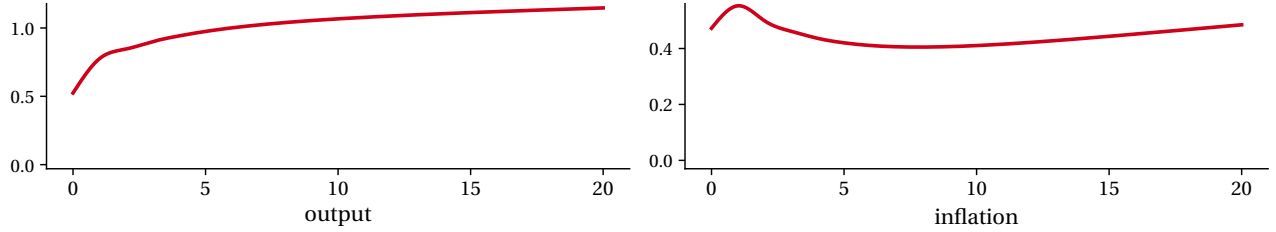


Figure 1. The Power of Forward Guidance

The key to understanding Figure 1 is that agents use a statistical model that provides a good in-sample fit to their past observations to make an out-of-sample forecast in a new regime with forward guidance. The observations they use to estimate their model are from a period without forward guidance, so the statistical relationships among variables are different than they would be under forward guidance. For instance, under the agents' pseudo-true one-state model, the correlation of \hat{i}_t and \hat{x}_{t-1} is only -0.17 , while the correlation of \hat{i}_t and $\hat{\pi}_{t-1}$ is about 0.73 .³¹ More generally, agents perceive the correlation of \hat{i}_t and \hat{x}_{t-s} to be -0.17×0.985^s and the correlation of \hat{i}_t and $\hat{\pi}_{t-s}$ to be 0.74×0.985^s . Therefore, changes in the interest rate in period $t + T$ have a modest impact on the agents' time- t forecasts of the output gap and inflation rate in periods $t + 1, t + 2, \dots, t + T - 1$, and the impact decays exponentially with T .

³¹The true correlation of \hat{i}_t and \hat{x}_{t-1} in the data used to estimate the model is only about -0.11 , whereas the true correlation of \hat{i}_t and $\hat{\pi}_{t-1}$ is 0.71 . Agents overestimate the extent to which changes in the next period's interest rate affect the current output gap and inflation rate, but not by much.

6 Application to the Real Business Cycle Model

For my second application, I consider the textbook real business cycle (RBC) model.

6.1 Primitives

Preferences, technology, and market structure are standard. Households value consumption and labor according to the per-period utility function

$$u(c, n) = \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \psi \frac{n^{1+\varphi}}{1+\varphi},$$

where c denotes consumption, n denotes labor, σ is the elasticity of intertemporal substitution (EIS), φ is the inverse Frisch elasticity of labor supply, and ψ is a constant that determines the steady-state working hours. The consumption good is produced by a measure of competitive firms by combining labor and capital according to the Cobb–Douglas production function

$$o_t = a_t k_t^\alpha n_t^{1-\alpha},$$

where o_t denotes output, a_t is total-factor productivity (TFP), and k_t denotes the capital stock at the beginning of period t . TFP follows a first-order autoregressive process in logs: $\hat{a}_t \equiv \log a_t$, and

$$\hat{a}_t = \rho \hat{a}_{t-1} + \epsilon_t. \quad (17)$$

In every period, households choose consumption, labor supply, and the next period's capital stock subject to the following flow budget constraint:

$$k_{t+1} = (1 - \delta + r_t)k_t + w_t n_t - c_t,$$

where δ denotes the depreciation rate of capital, r_t is the rental rate of capital, and w_t is the wage rate. Finally, market clearing determines investment:

$$i_t = o_t - c_t.$$

6.2 Log-linear Temporary Equilibrium

As is common in the literature, I log-linearize the model around a steady state in which $\hat{a}_t = 0$, $o_t = o$, $w_t = w$, $r_t = r$, $n_t = n$, $i_t = i$, $k_t = k$, and $c_t = c$. The usual aggregate Euler equation may not hold away from rational expectations. I instead start by characterizing the temporary equilibrium relations, which impose individual optimality and market clearing conditions but not rational expectations. The log-linearized temporary-equilibrium conditions are given by

$$\hat{o}_t = \hat{a}_t + \alpha \hat{k}_t + (1 - \alpha) \hat{n}_t, \quad (18)$$

$$\hat{w}_t = \hat{a}_t + \alpha(\hat{k}_t - \hat{n}_t), \quad (19)$$

$$\hat{r}_t = r \hat{a}_t + (1 - \alpha)r(\hat{n}_t - \hat{k}_t), \quad (20)$$

$$\hat{n}_t = \frac{1}{\varphi} \hat{w}_t - \frac{1}{\sigma \varphi} \hat{c}_t, \quad (21)$$

$$\hat{i}_t = \frac{o}{i} \hat{o}_t - \frac{c}{i} \hat{c}_t, \quad (22)$$

$$\hat{k}_t = (1 - \delta) \hat{k}_{t-1} + \delta \hat{i}_{t-1}, \quad (23)$$

$$\hat{c}_t = \frac{\chi}{\beta} \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + (\chi - \beta \sigma) \sum_{s=1}^{\infty} \beta^s E_t[\hat{r}_{t+s}] + \chi \zeta \sum_{s=1}^{\infty} \beta^s E_t[\hat{w}_{t+s}], \quad (24)$$

where

$$\chi \equiv (1 - \beta) \left(\frac{(1 - \alpha)r}{\alpha \sigma \varphi} + \frac{c}{k} \right)^{-1},$$

$$\zeta \equiv \frac{(1 - \alpha)(1 + \varphi)r}{\alpha \varphi},$$

and $E_t[\cdot]$ denotes the subjective forecast of households. The details of this derivation can be found in Online Appendix E. Equations (17)–(24) fully characterize the equilibrium of the economy once the subjective expectations are specified.

6.3 Subjective Expectations and Equilibrium

I find the equilibrium both under rational expectations and assuming that households are constrained to d -state models. In both cases, I assume that households know the steady-state values of every variable and perfectly observe the vector $y_t \equiv (\hat{a}_t, \hat{o}_t, \hat{w}_t, \hat{r}_t, \hat{n}_t, \hat{i}_t, \hat{k}_t, \hat{c}_t)$ of log-deviations from the steady state.

Where the two cases differ is in how households model log-deviations from the steady state. Under RE, the households' model of y_t coincides with the observable's true, equilibrium stochastic process. When households are constrained to d -state models, they believe that y_t follows a d -dimensional model of the form (1) and use a pseudo-true d -state model to forecast the future values of y .

The equilibrium definition when households use pseudo-true d -state models is straightforward. An equilibrium consists of a stochastic process \mathbb{P} for y_t and a model θ^* for households such that (i) \mathbb{P} is derived from market-clearing conditions and households' optimal consumption, labor supply, and investment behavior given their subjective model θ^* , and (ii) θ^* is a pseudo-true d -state model given the stochastic process \mathbb{P} .

The rational expectations equilibrium has a two-state representation. Therefore, agents constrained to d -state models with $d > 1$ recover the true process, and the equilibrium given d -state models with $d > 1$ coincides with the rational expectations equilibrium. This observation highlights the fact that constraining agents to d -state models represents the only deviation from the full-information rational-expectations benchmark. Furthermore, the constraint is slack as long as $d > 1$.

Finding an equilibrium involves solving a fixed-point equation. The rational expectations equilibrium can be found using existing techniques. In Online Appendix E, I discuss how one

can find the constrained rational expectations equilibrium in the case where households use pseudo-true d -state models with $d = 1$. In the same appendix, I also provide a more formal definition of equilibrium.

6.4 Calibration

The exogenous parameters of the model are calibrated as follows: A period represents a quarter. The quarterly discount rate is set to $\beta = 0.99$. The EIS and the Frisch elasticity of labor supply are both set to one. The depreciation rate is set to $\delta = 0.012$ and the capital share of output to $\alpha = 0.3$. TFP has a persistence parameter of $\rho = 0.95$. I set the standard deviation of TFP innovations to one.

Note that d is the only extra free parameter relative to the full-information rational-expectations version of the model. Once one chooses a value for d , the expectations are fully pinned down by the primitives of the economy (as in the benchmark). Moreover, the d -state equilibrium nests the RE equilibrium by setting $d > 1$.

6.5 Results

Figure 2 plots the impulse-response functions to a one percent increase in TFP. The responses in the rational expectations case are in dashed green, whereas those for the case in which agents use pseudo-true one-state models are in solid red. Every variable except for the rental rate of capital is measured in log changes from its steady state value; the rental rate of capital is measured in percentage point changes from its steady state value. The variable labeled the “state of the economy” is defined as the households’ nowcast \hat{z}_t of the subjective state z_t in their subjective model of the economy. Since the scale of z_t is not identifiable either to the agents or the econometrician, the scale of \hat{z}_t is intrinsically meaningless.

The state of the economy at time t can be expressed as a linear combination of the time- t values of the capital stock and TFP, with the weights determined endogenously in equilibrium:³²

$$\hat{z}_t = 0.947\hat{k}_t + 0.053\hat{a}_t.$$

The state of the economy is much more sensitive to changes in the capital stock than to changes in TFP. This can also be seen in the impulse-response functions: The state of the economy co-moves almost perfectly with capital.

The fact that the state of the economy inherits the dynamics of the capital stock is a manifestation of persistence bias. In equilibrium, the capital stock is more persistent than TFP, as can be seen from the impulse-response functions. Therefore, the households’ nowcast of the subjective state moves almost one-for-one with changes in the capital stock.

Consumption, in turn, inherits the dynamics of the state of the economy. Since β is close to one in my calibration, consumption is almost purely forward looking. Therefore, it moves in

³²As previously mentioned, the magnitude of \hat{z}_t is irrelevant. I normalize \hat{z}_t so that $\hat{z}_t = p_k\hat{k}_t + p_a\hat{a}_t$ with $|p_k| + |p_a| = 1$.

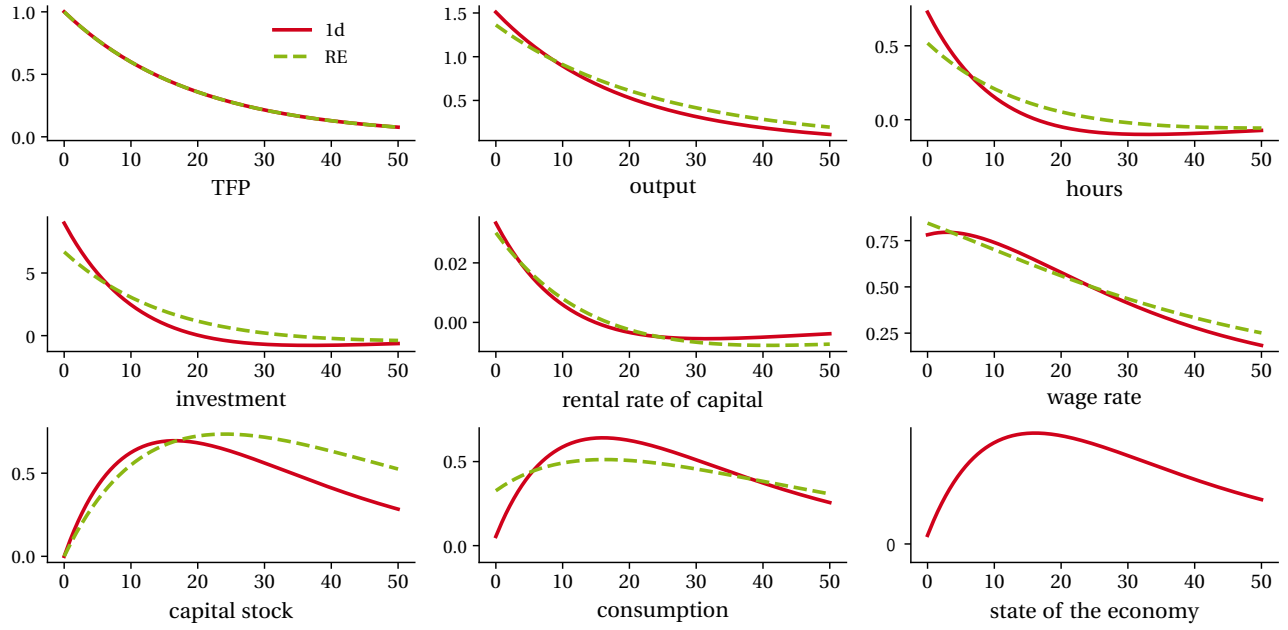


Figure 2. Impulse Response Functions to a TFP Shock

tandem with changes in the households' forecasts, which in turn, move almost one-for-one with their nowcast \hat{z}_t of the subjective state. In equilibrium,

$$\hat{c}_t = 0.089\hat{k}_t + 0.088\hat{r}_t + 0.0091\hat{w}_t + 0.841\hat{z}_t.$$

That is, consumption is much more sensitive to changes in the state of the economy than to current prices and quantities.

The upshot is that consumption comoves with the capital stock. The (unconditional) correlation of consumption and the capital stock is 0.999 when households use pseudo-true one-state models; in contrast, it is 0.956 when households have rational expectations. Even though consumption is an almost purely forward-looking variable, it is anchored to the most persistent backward-looking variable in the economy: capital.

The fact that consumption is anchored to capital dampens the initial response of consumption to TFP shocks. The response of consumption on impact when households use a pseudo-true one-state model is 83% smaller than the corresponding response under RE. The consumption response in the one-state case continues to be smaller than the RE response for six quarters after impact. But as the one-state economy builds up its capital stock, the households' view of the state of the economy improves and their consumption increases. At some point, consumption in the one-state economy overshoots its RE counterpart. The model thus provides a parsimonious account of the hump-shaped response of consumption to TFP in empirical studies.^{33 34}

The initial underreaction and the subsequent overshooting of consumption increases its un-

³³For a meta-analysis of the response of aggregate variables to technology shocks, see Ramey (2016, pp. 135–151).

³⁴The response of consumption to TFP shocks is hump-shaped in the model even when TFP is i.i.d. over time, suggesting a resolution to Cogley and Nason (1993)'s observation that the RBC model has a weak propagation mechanism.

conditional variance by about 10% relative to the RE benchmark. Similarly, hours and investment are about 41% and 24% more volatile, respectively, when agents are constrained to one-state models relative to the RE benchmark. The increased volatility of consumption and hours, resulting from the agents' use of simple models, leads to an increase in the cost of business cycles.

7 Application to the Diamond–Mortensen–Pissarides Model

As my last application of the framework, I study how the predictions of the standard labor search and matching model change when agents are constrained to use simple models. I do so in the context of the stochastic version of the Diamond–Mortensen–Pissarides (DMP) model in discrete time. I start by describing the primitives of the economy.

7.1 Primitives

There is a continuum of workers and firms in the economy. The mass of workers is normalized to one, whereas the mass of firms is determined by free entry. Workers and firms are both risk neutral and discount the future at rate β . A worker matched with a firm generates a_t units of output in each period, whereas an unemployed worker produces $b < 1$ units. I assume that $a_t - b = (1 - b) \exp(\hat{a}_t)$, where \hat{a}_t is a shock to labor productivity net of home production. This specification of labor productivity guarantees that $a_t > b$ for all t , so it is always efficient for workers to be employed by firms.

Unemployed workers and firms randomly match in a frictional labor market. A matching function determines the rate at which unemployed workers meet firms. Each unemployed worker finds a job in period t with probability $p_t = \mu \theta_t^{1-\alpha}$, and each vacancy is filled with probability $q_t = \mu \theta_t^{-\alpha}$, where $\theta_t \equiv v_t/u_t$ denotes market tightness, i.e., the ratio of the number of vacancies to the unemployment rate, and μ and α are parameters of the matching function. Each job is destroyed in each period with probability $s_t = s \exp(\hat{s}_t)$, where \hat{s}_t is a separation shock. Firms incur a cost k per period (measured in units of output) for maintaining a vacancy.

Wages are determined through Nash bargaining between a worker and a firm, with the threat point of the worker the value of unemployment, the threat point of the firm the value of an unfilled vacancy (which will be zero in equilibrium), and the worker's bargaining power equal to δ .

I assume that net labor productivity and separation rate shocks follow the autoregressive process

$$\begin{pmatrix} \hat{a}_t \\ \hat{s}_t \end{pmatrix} = \begin{pmatrix} \rho_a & 0 \\ 0 & \rho_s \end{pmatrix} \begin{pmatrix} \hat{a}_{t-1} \\ \hat{s}_{t-1} \end{pmatrix} + \epsilon_t, \quad (25)$$

where $\epsilon_t \sim \mathcal{N}(0, \Sigma)$. This specification allows for labor productivity and separation rate to be correlated, as is the case in the data.

7.2 Temporary Equilibrium

The recursive equations that characterize the solution to the DMP model may not hold without rational expectations. I instead start by characterizing the temporary-equilibrium relations, which hold under arbitrary expectations. I assume that firms and workers use state-space models with the same number of states, ending up with the same subjective expectations in equilibrium. Market tightness and wage then satisfy the following equations:³⁵

$$\theta_t^\alpha = \frac{\mu}{k} E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau \prod_{k=1}^{\tau-1} (1 - s_{t+k}) (a_{t+\tau} - w_{t+\tau}) \right], \quad (26)$$

$$w_t = \delta a_t + (1 - \delta)b + \delta E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau \prod_{k=0}^{\tau-1} (1 - s_{t+k}) (a_{t+\tau} - w_{t+\tau}) \right] - (1 - \delta) E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau \prod_{k=0}^{\tau-1} (1 - s_{t+k} - p_{t+k}) (w_{t+\tau} - b) \right], \quad (27)$$

where I use the convention $\prod_{k=1}^0 (1 - s_{t+k}) (a_{t+\tau} - w_{t+\tau}) = 1$ in equation (26). The unemployment rate follows the first-order difference equation

$$u_t = u_{t-1} + s_{t-1}(1 - u_{t-1}) - \mu \theta_{t-1}^{1-\alpha} u_{t-1}. \quad (28)$$

Equations (25)–(28) together with the specification of the subjective expectations fully characterize the equilibrium. The derivation of these equations and other omitted calculations from this section can be found in Online Appendix F. To simplify the numerical computations, I log-linearize the temporary equilibrium of the economy around a steady state in which $a_t = 1 > b$ and $s_t = s$.

7.3 Subjective Expectations and Equilibrium

I solve the model both under rational expectations and assuming that agents are constrained to one-state models. In both cases, I assume that every agent knows the steady-state value of every variable and perfectly observes vector $y_t \equiv (\hat{a}_t, \hat{s}_t, \hat{\theta}_t, \hat{v}_t, \hat{u}_t, \hat{p}_t, \hat{q}_t, \hat{w}_t)$ of log-deviations from the steady state. Agents constrained to one-state models believe that y_t follows a one-state model of the form (1) for some $\theta = (A, B, Q, R)$ and use a pseudo-true one-state model to forecast the future values of y .

Equilibrium is defined as in the previous applications. It consists of a stochastic process \mathbb{P} for y_t and a model θ^* for the agents such that (i) \mathbb{P} is derived from the agents' optimal behavior given their subjective model θ^* , and (ii) θ^* is a pseudo-true one-state model given the stochastic process \mathbb{P} . A more formal definition can be found in Online Appendix F.

³⁵Nash bargaining only determines the total value delivered to workers and firms and not the timing of the payoffs or the wage rate. To determine the wage, I assume that workers and firms both take the future expected wages as given and adjust the current wage to split the surplus according to the Nash bargaining solution.

7.4 Calibration

The model is calibrated as follows: Each period corresponds to a month. The discount factor is set to $\beta = 0.99$. I set the mean of the separation rate to $s = 0.035$, so jobs last for about 2.5 years on average. The steady-state job-finding probability is set to $p = 0.4$ per month. The elasticity parameter of the matching function is set to $\alpha = 0.72$. The workers' bargaining power is set to the same value: $\delta = 0.72$. Setting $\delta = \alpha$ ensures that the Hosios condition is satisfied. I set the persistence parameter of the shock to $\rho_a = 0.96$ for labor productivity and $\rho_s = 0.90$ for separation rate. I normalize the steady-state output per worker to $a = 1$. The flow payoff to workers from unemployment is set to $b = 0.4$.³⁶

The impulse-response functions are independent of the volatility of the shocks and their correlation when agents have rational expectations—but not when they are constrained to one-dimensional models. I set the correlation of labor productivity and separation rate shocks to -0.4 and the ratio of the standard deviation of labor productivity to that of separation rate to ten. These choices ensure that the (pairwise) correlation coefficients between labor productivity, separation rate, and the unemployment rate are broadly consistent with the data in [Shimer \(2005\)](#). Finally, I normalize the standard deviation of labor productivity to one. The results that follow do not depend on this normalization.

7.5 Results

Figures 3 and 4 plot the impulse-response functions to a one percent increase in labor productivity and separation rate, respectively. The responses in the rational expectations case are in dashed green, whereas those for the case in which agents use pseudo-true one-state models are in solid red. Variables are all measured in log changes from their steady state values. As in the previous application, the variable labeled the “state of the economy” is defined as the agents' nowcast \hat{z}_t of the subjective state z_t . Since the scale of z_t is not identifiable either to the agents or the econometrician, the scale of \hat{z}_t is intrinsically meaningless. However, the two panels plotting the response of \hat{z}_t to labor productivity and separation rate shocks use the same scale, so the responses of the state of the economy are comparable across the two shocks.

The state of the economy at time t can be expressed as a linear combination of the time- t values of the unemployment rate, labor productivity, and separation rate with the weights determined endogenously in equilibrium:³⁷

$$\hat{z}_t = -0.812\hat{u}_t + 0.010\hat{a}_t - 0.177\hat{s}_t. \quad (29)$$

The state of the economy is almost five times more sensitive to changes in the unemployment rate than to changes in separation rate, and it barely responds to changes in labor productivity.

³⁶These parameter values are all consistent with the calibration in [Shimer \(2005\)](#). Others, such as [Hagedorn and Manovskii \(2008\)](#), rely on values of b closer to one to amplify the response of unemployment to labor productivity shocks.

³⁷I normalize \hat{z}_t so that $\hat{z}_t = p_u\hat{u}_t + p_a\hat{a}_t + p_s\hat{s}_t$ with $|p_u| + |p_a| + |p_s| = 1$.

Since the shocks are correlated with each other, the mapping from the persistence of the shocks to their weights in the expression for \hat{z}_t is not as simple as in the RBC application. Rather, the relative attention vector p , which determines the weights of different variables in the determination of the state of economy, depends on the joint dynamics of the unemployment rate, labor productivity, and the separation rate as in Theorem 5. Those dynamics, in turn, are determined in equilibrium as a function, among other things, of the attention vector p .

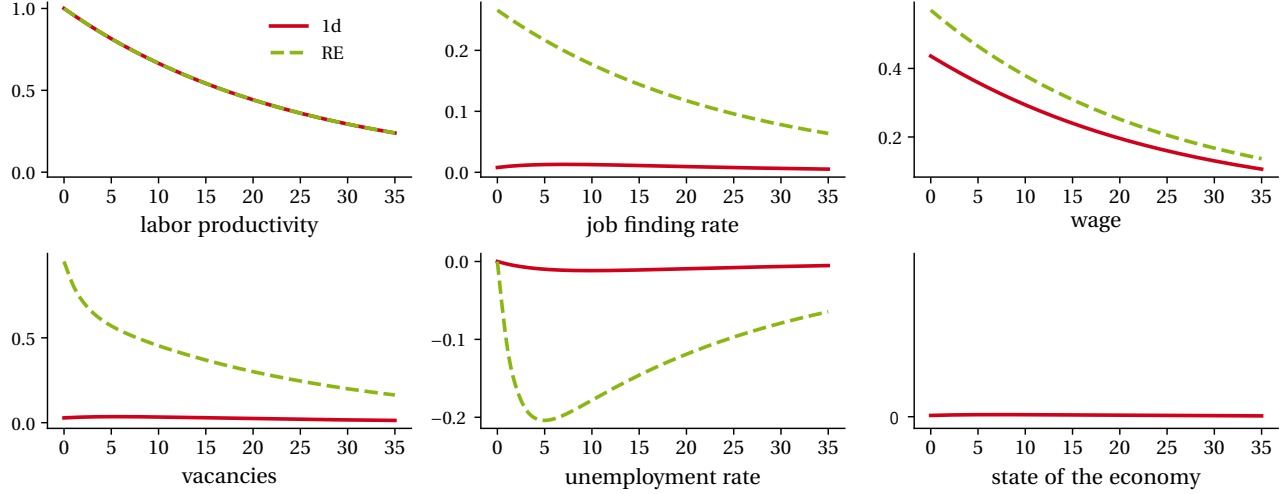


Figure 3. Impulse Response Functions to a Labor Productivity Shock

The key to understanding the economy's response to labor productivity shocks is the observation that current productivity has no direct effect on the agents' decisions. The only interesting decision in the DMP model is the firms' vacancy-creation decision, which together with the current unemployment rate, fully determines market tightness through equation (26). Market tightness and the exogenous separation rate, in turn, determine next period's unemployment rate through equation (28). Current productivity appears nowhere in these equations; it is only the firms' forecasts of future productivity that enters the dynamics of market tightness, job-finding rate, and unemployment. In fact, in equilibrium,

$$\begin{aligned}\hat{\theta}_t &= 2.76\hat{z}_t, \\ \hat{p}_t &= 0.774\hat{z}_t.\end{aligned}$$

That is, market tightness and job-finding rate are perfectly correlated with the state of the economy.

This property of the DMP model leads to a form of complementarity in the agents' relative attention to different observables, which, in equilibrium, dampens the response of the economy to labor productivity shocks. Suppose firms reduce the weight they assign to labor productivity when forming their estimate of the current state of the economy. Doing so makes their forecasts less sensitive to current labor productivity and dampens the effect of labor productivity on unemployment fluctuations. This, in turn, reduces labor productivity's weight in the agents'

estimate of the state of the economy under a pseudo-true model. In equilibrium, labor productivity receives little weight in the agents' forecasts and has a small impact on the dynamics of vacancies, job-finding rate, and unemployment.³⁸

The economy's response to a separation rate shock is perhaps even more subtle. Under rational expectations, an increase in separation rate foreshadows an increase in the unemployment rate. This increase in the unemployment rate will be beneficial to would-be employers: A higher unemployment rate means a slacker labor market and a higher job-filling rate. This makes it more likely that a firm will recoup the cost of creating a vacancy, thus leading to an increase in the number of vacancies through the free-entry condition. This dynamic is behind the counterfactual positive correlation between the number of vacancies and the unemployment rate in a DMP model with only separation rate shocks.

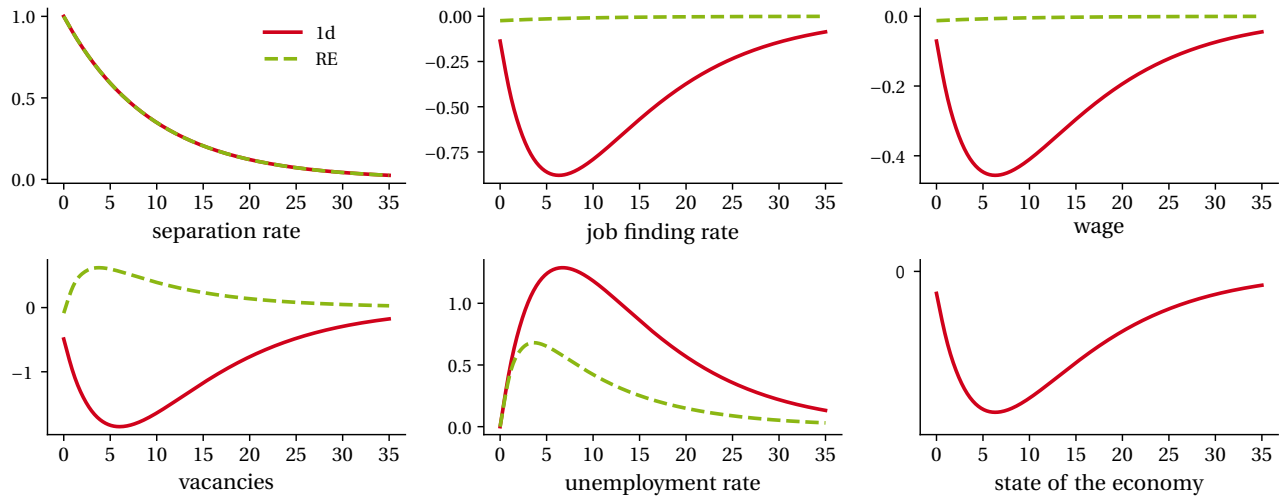


Figure 4. Impulse Response Functions to a Separation Rate Shock

Constraining agents to low-dimensional models turns this dynamic on its head. By equation (29), an increase in separations lowers the agents' nowcast of the state of the economy both directly and indirectly through the resulting increase in unemployment. The deterioration in the firms' nowcast lowers their expectations of returns to posting a vacancy. This decrease is large enough (at least, in the current calibration) to overturn the effect of the increase in the job-filling rate. As a result, firms post fewer vacancies, causing an even bigger increase in the unemployment rate.

The recession that follows an increase in separations in the one-state version of the model has a Keynesian flavor. The increase in separations and the resulting increase in unemployment make firms pessimistic. They respond by slowing their recruiting activities, which, in turn, exacerbates the unemployment problem, darkening the outlook further, and so on. The result is an inefficiently deep and long recession.

³⁸Note that the wages rate is highly sensitive to labor productivity even when agents rely on pseudo-true one-state models. This is due to the fact that labor productivity has a direct effect on the current wage, as can be seen in equation (27).

The inability of the standard DMP model to generate realistic unemployment fluctuations in response to realistic productivity and separation shocks is known as the [Shimer \(2005\)](#) puzzle. This puzzle has led to a large literature, which aims to resolve it by modifying the DMP model or Shimer’s calibration of it.

The exercise in this section suggests a novel path forward. It shows that constraining agents to use simple models allows even the most basic DMP model to exhibit significant amplification, propagation, and comovement in response to separation shocks, bringing its behavior more in line with the stylized facts about employment fluctuations.

8 Concluding Remarks

This paper suggests a novel approach to modeling bounded rationality that is portable across different applications. I illustrated the use of the framework in three canonical workhorse macroeconomic models. I showed that constraining agents to use estimated simple models does not alter the divine coincidence property of the standard new-Keynesian model but significantly reduces the power of forward guidance. Characterizing the optimal monetary policy in a new-Keynesian model with simple agents is a promising direction for future research.

I intentionally focused on bare-bones macro models to allow for a transparent discussion of how simple models work. However, simple models can be easily embedded in modern heterogeneous-agent macro models to study the implications of bounded rationality. One can do so because neither the additional degrees of freedom nor the computational complexity of finding pseudo-true models scale with the size of the macro model. This paper’s approach can also be extended to allow for heterogeneity in d without having to contend with the complications associated with heterogeneous-belief macro models (such as the “infinite regress” problem).

The characterization results of the paper primarily focused on the complete information benchmark, where agents’ forecasts end up being Markovian in observables. In reality, economic agents face both incomplete information and limits to their models’ complexity. Combining the two leads to richer models of expectations formation that exhibit features of incomplete information models (such as history dependence and slow adjustment) and simple models (such as persistence bias and comovement).

Throughout the paper, I took dimension d of the agents’ model as a fixed parameter. This parameter can be identified using expectations data. It can also be estimated jointly with other parameters in DSGE models only using data on aggregate variables. I leave the problem of estimating d to future research.

Appendices

A Weighted Mean-Squared Forecast Error

The agent's time- t one-step-ahead forecast error given model θ is defined as

$$e_t(\theta) \equiv y_{t+1} - E_t^\theta[y_{t+1}],$$

where E_t^θ denotes the agent's subjective expectation conditional on her information at time t and given model θ . The weighted average of mean-squared forecast errors given a symmetric weight matrix $W \in \mathbb{R}^{n \times n}$ is defined as

$$\text{MSE}_W(\theta) = \mathbb{E} \left[e_t'(\theta) W e_t(\theta) \right].$$

Instead of assuming that the agent uses a model that minimizes the KLDR, one can assume that she makes their forecasts using a model θ that minimizes $\text{MSE}_W(\theta)$ for some matrix W .

Using the mean-squared forecast error as the notion of fit has two disadvantages relative to the KLDR. First, the choice of matrix W introduces additional degrees of freedom when the observable is not a scalar. Second, the minimizer of weighted mean-squared errors is in general not invariant to linear transformations of the vector of observable (unless if the weight matrix W is transformed accordingly).

Let θ^* denote a pseudo-true d -state model, and let $\hat{\Sigma}_y^{d*}$ denote the implied subjective variance of y_{t+1} conditional on the agent's information at time t .

Proposition A.1. *If W is equal to the inverse of $\hat{\Sigma}_y^{d*}$, then $\theta^* \in \arg \min_{\theta \in \Theta_d} \text{MSE}_W(\theta)$.*

The proposition establishes that mean-squared forecast error minimization coincides with KLDR minimization under the appropriate choice of the weighting matrix W .

B Partial Equilibrium and General Equilibrium

In this appendix, I argue that the implications of the general framework are largely unchanged in a general equilibrium setting where the observables' laws of motion depend on the agents' actions. I consider a stylized general equilibrium (GE) economy in which observables are linear functions of some exogenous shocks and agents' actions. Specifically, I assume that, in equilibrium, the vector of observables $y_t \in \mathbb{R}^n$ can be written as

$$y_t^{\text{GE}} = \tilde{H}' f_t + g x_t^{\text{GE}}, \quad (\text{B.1})$$

where $x_t \in \mathbb{R}$ is the agents' action, $f_t \in \mathbb{R}^m$ is the vector of exogenous shocks, $\tilde{H} \in \mathbb{R}^{m \times n}$ is a rank- m matrix, and $g \in \mathbb{R}^n$ is a vector that parameterizes the strength of the GE feedback from agents' actions to the aggregate observable. The agents' best-response functions are given by

$$x_t^{\text{GE}} = b' y_t^{\text{GE}} + E_t \left[\sum_{s=1}^{\infty} \beta^s c' y_{t+s}^{\text{GE}} \right]. \quad (\text{B.2})$$

For simplicity, I assume that the shocks follow m independent AR(1) processes:

$$f_t = Ff_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma), \quad (\text{B.3})$$

where $F = \text{diag}(\alpha_1, \dots, \alpha_m)$ and $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$. Equations (B.1)–(B.3) together with the specification of agents' subjective expectations fully characterize the (general) equilibrium of the economy.

I contrast this economy with a partial equilibrium (PE) economy in which

$$y_t^{\text{PE}} = H' f_t, \quad (\text{B.4})$$

$$x_t^{\text{PE}} = b' y_t^{\text{PE}} + E_t \left[\sum_{s=1}^{\infty} \beta^s c' y_{t+s}^{\text{PE}} \right], \quad (\text{B.5})$$

and f_t follows (B.3).

The term “partial equilibrium” is inspired by the following hypothetical scenario: Suppose we considered the economy described by equations (B.1)–(B.3) but ignored the fact that agents' actions affect the observable, which in turn affect agents' actions, and so on. Then the response of the GE economy to shocks would be described by equations (B.4)–(B.5). The following result establishes an observational equivalence between the GE and PE economies:

Proposition B.1. *Consider the general equilibrium economy (B.1)–(B.3) and the partial equilibrium economy (B.3)–(B.5), and suppose that, in each economy, agents use pseudo-true m.i.o. d -state models to forecast the observable. If*

$$\tilde{H} = H \left(I - \left(b + \sum_{k=1}^d \frac{\alpha_k \beta}{1 - \alpha_k \beta} H^\dagger e_k e_k' H c \right) g' \right),$$

then the linear equilibria of the two economies are observationally equivalent.

Several remarks are in order. First, the result is a corollary of the linear-invariance result (Theorem 2) and the fact that agents' actions are linear in the observable. Second, the proposition covers the rational-expectations case by setting $d = m$. Third, when $\beta = 0$, the effect of going from PE to GE is to amplify the response of observables to shocks, as measured by matrix H' , by the GE multiplier $(I - gb')^{-1}$. When $\beta > 0$, the multiplier has an additional term, which captures the general equilibrium effect of the updating of expectations by agents.

Last but not least, the distinctions between exogenous and endogenous variables, on one hand, and PE and GE, on the other, are largely inconsequential in this framework. Agents' expectations of endogenous variables are consistent with their expectations of exogenous variables and the structural equations of the economy, the GE economy is just the PE economy with a linearly transformed H matrix, and agents' expectations in the GE economy are just linear transformations of their expectations in the PE economy.

References

- Adam, Klaus and Albert Marcet (2011), “Internal Rationality, Imperfect Market Knowledge and Asset Prices.” *Journal of Economic Theory*, 146, 1224–1252.
- Adam, Klaus, Albert Marcet, and Johannes Beutel (2017), “Stock Price Booms and Expected Capital Gains.” *American Economic Review*, 107, 2352–2408.
- Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2021), “Overreaction in Expectations: Evidence and Theory.” Working paper.
- Akaike, Hirotugu (1975), “Markovian Representation of Stochastic Processes by Canonical Variables.” *SIAM Journal on Control*, 13, 162–173.
- Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello (2015), “Monetary Shocks in Models with Inattentive Producers.” *The Review of Economic Studies*, 83, 421–459.
- Anderson, Brian D.O. and John B. Moore (2005), *Optimal Filtering*. Dover Publications, Mineola, N.Y.
- Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2020), “Business-Cycle Anatomy.” *American Economic Review*, 110, 3030–70.
- Angeletos, George-Marios and Zhen Huo (2021), “Myopia and Anchoring.” *American Economic Review*, 111, 1166–1200.
- Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2021), “Imperfect Macroeconomic Expectations: Evidence and Theory.” *NBER Macroeconomics Annual*, 35, 1–86.
- Angeletos, George-Marios and Jennifer La’O (2009), “Incomplete Information, Higher-Order Beliefs and Price Inertia.” *Journal of Monetary Economics*, 56, S19–S37.
- Angeletos, George-Marios and Chen Lian (2018), “Forward Guidance without Common Knowledge.” *American Economic Review*, 108, 2477–2512.
- Berk, Robert H. (1966), “Limiting Behavior of Posterior Distributions When the Model is Incorrect.” *Annals of Mathematical Statistics*, 37, 51–58.
- Bianchi, Francesco, Cosmin L. Ilut, and Hikaru Saijo (2021), “Diagnostic Business Cycles.” Working paper.
- Bidder, Rhys and Ian Dew-Becker (2016), “Long-Run Risk Is the Worst-Case Scenario.” *American Economic Review*, 106, 2494–2527.
- Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (2020), “Overreaction in Macroeconomic Expectations.” *American Economic Review*, 110, 2748–82.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018), “Diagnostic Expectations and Credit Cycles.” *Journal of Finance*, 73, 199–227.
- Bray, M. M. and N. E. Savin (1986), “Rational Expectations Equilibria, Learning, and Model Specification.” *Econometrica*, 54, 1129–1160.
- Bray, Margaret (1982), “Learning, Estimation, and the Stability of Rational Expectations.” *Journal of Economic Theory*, 26, 318–339.
- Broer, Tobias and Alexandre N. Kohlhas (2020), “Forecaster (Mis-) Behavior.” Working paper.
- Bullard, James, George W. Evans, and Seppo Honkapohja (2008), “Monetary Policy, Judgment, and Near-Rational Exuberance.” *American Economic Review*, 98, 1163–1177.
- Bullard, James and Kaushik Mitra (2002), “Learning About Monetary Policy Rules.” *Journal of Monetary Economics*, 49, 1105–1129.

- Bunke, Olaf and Xavier Milhaud (1998), “Asymptotic Behavior of Bayes Estimates Under Possibly Incorrect Models.” *Annals of Statistics*, 26, 617–644.
- Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner (2021), “Sectoral Media Focus and Aggregate Fluctuations.” *American Economic Review*, 111, 3872–3922.
- Cho, In Koo, Noah Williams, and Thomas J. Sargent (2002), “Escaping Nash Inflation.” *Review of Economic Studies*, 69, 1–40.
- Cogley, Timothy and James M. Nason (1993), “Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model.” *Economics Letters*, 43, 77–81.
- Coibion, Olivier and Yuriy Gorodnichenko (2015), “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts.” *American Economic Review*, 105, 2644–78.
- Dew-Becker, Ian and Charles G. Nathanson (2019), “Directed Attention and Nonparametric Learning.” *Journal of Economic Theory*, 181, 461–496.
- Douc, Randal and Eric Moulines (2012), “Asymptotic Properties of the Maximum Likelihood Estimation in Misspecified Hidden Markov Models.” *Annals of Statistics*, 40, 2697–2732.
- Esponda, Ignacio and Demian Pouzo (2016), “Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models.” *Econometrica*, 84, 1093–1130.
- Esponda, Ignacio and Demian Pouzo (2021), “Equilibrium in Misspecified Markov Decision Processes.” *Theoretical Economics*, 16, 717–757.
- Eusepi, Stefano and Bruce Preston (2011), “Expectations, Learning, and Business Cycle Fluctuations.” *American Economic Review*, 101, 2844–2872.
- Eusepi, Stefano and Bruce Preston (2018), “The Science of Monetary Policy: An Imperfect Knowledge Perspective.” *Journal of Economic Literature*, 56, 3–59.
- Evans, George W. and Seppo Honkapohja (1995), “Local Convergence of Recursive Learning to Steady States and Cycles in Stochastic Nonlinear Models.” *Econometrica*, 63, 195–206.
- Evans, George W. and Seppo Honkapohja (2012), *Learning and Expectations in Macroeconomics*. Princeton University Press.
- Farhi, Emmanuel and Iván Werning (2019), “Monetary Policy, Bounded Rationality, and Incomplete Markets.” *American Economic Review*, 109, 3887–3928.
- Faure, Pierre L. (1976), “Stochastic Realization Algorithms.” In *Mathematics in Science and Engineering*, volume 126, 1–25, Elsevier.
- Forni, Mario and Marco Lippi (2001), “The Generalized Dynamic Factor Model: Representation Theory.” *Econometric Theory*, 17, 1113–1141.
- Fuster, Andreas, Benjamin Hebert, and David Laibson (2012), “Investment Dynamics with Natural Expectations.” *International Journal of Central Banking*, 8, 243–265.
- Fuster, Andreas, David Laibson, and Brock Mendel (2010), “Natural Expectations and Macroeconomic Fluctuations.” *Journal of Economic Perspectives*, 24, 67–84.
- Gabaix, Xavier (2014), “A Sparsity-Based Model of Bounded Rationality.” *Quarterly Journal of Economics*, 129, 1661–1710.
- Gabaix, Xavier (2020), “A Behavioral New Keynesian Model.” *American Economic Review*, 110, 2271–2327. Working paper.

- Galí, Jordi (2015), *Monetary Policy, Inflation, and the Business Cycle*, second edition. Princeton University Press.
- García-Schmidt, Mariana and Michael Woodford (2019), “Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis.” *American Economic Review*, 109, 86–120.
- Gáti, Laura (2020), “Monetary Policy & Anchored Expectations: An Endogenous Gain Learning Model.” Working Paper.
- Gerko, Elena (2017), “Expectations and Monetary Policy.” Working paper.
- Gevers, Michel and Vincent Wertz (1984), “Uniquely Identifiable State-Space and ARMA Parametrizations for Multivariable Linear Systems.” *Automatica*, 20, 333–347.
- Grandmont, Jean-Michel (1977), “Temporary General Equilibrium Theory.” *Econometrica*, 45, 535–572.
- Hagedorn, Marcus and Iouri Manovskii (2008), “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited.” *American Economic Review*, 98, 1692–1706.
- Hansen, Lars Peter and Thomas J. Sargent (2008), *Robustness*. Princeton University Press.
- Ho, B. L. and R. E. Kálmán (1966), “Effective Construction of Linear State-Variable Models from Input/Output Functions.” *at-Automatisierungstechnik*, 14, 545–548.
- Katayama, Tohru (2005), *Subspace Methods for System Identification*, volume 1. Springer.
- Kleijn, B. J. K. and A. W. Van Der Vaart (2006), “Misspecification in Infinite-Dimensional Bayesian Statistics.” *Annals of Statistics*, 34, 837–877.
- Kozłowski, Julian, Laura Veldkamp, and Venky Venkateswaran (2020), “The Tail That Wags the Economy: Beliefs and Persistent Stagnation.” *Journal of Political Economy*, 128, 2839–2879.
- Krusell, Per and Anthony A. Smith, Jr. (1998), “Income and Wealth Heterogeneity in the Macroeconomy.” *Journal of Political Economy*, 106, 867–896.
- Lorenzoni, Guido (2009), “A Theory of Demand Shocks.” *American Economic Review*, 99, 2050–84.
- Lucas, Robert E. (1972), “Expectations and the Neutrality of Money.” *Journal of Economic Theory*, 4, 103–124.
- Maćkowiak, Bartosz and Mirko Wiederholt (2009), “Optimal Sticky Prices under Rational Inattention.” *American Economic Review*, 99, 769–803.
- Maćkowiak, Bartosz and Mirko Wiederholt (2015), “Business Cycle Dynamics under Rational Inattention.” *Review of Economic Studies*, 82, 1502–1532.
- Mankiw, N. Gregory and Ricardo Reis (2002), “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.” *Quarterly Journal of Economics*, 117, 1295–1328.
- Marcet, Albert and Juan P. Nicolini (2003), “Recurrent Hyperinflations and Learning.” *American Economic Review*, 93, 1476–1498.
- Marcet, Albert and Thomas J. Sargent (1989), “Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information.” *Journal of Political Economy*, 97, 1306–1322.
- Meyn, Sean P. and Richard L. Tweedie (1993), *Markov Chains and Stochastic Stability*. Springer-Verlag.
- Milani, Fabio (2007), “Expectations, Learning and Macroeconomic Persistence.” *Journal of Monetary Economics*, 54, 2065–2082.

- Molavi, Pooya (2019), "Macroeconomics with Learning and Misspecification: A General Theory and Applications." Working paper.
- Molavi, Pooya, Alireza Tahbaz-Salehi, and Andrea Vedolin (2021), "Model Complexity, Expectations, and Asset Prices." Working Paper.
- Nimark, Kristoffer (2008), "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics*, 55, 365–382.
- Orphanides, Athanasios (2003), "Monetary Policy Evaluation With Noisy Information." *Journal of Monetary Economics*, 50, 605–631.
- Orphanides, Athanasios and John C. Williams (2005), "Inflation Scares and Forecast-Based Monetary Policy." *Review of Economic Dynamics*, 8, 498–527.
- Preston, Bruce (2005), "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter." *International Journal of Central Banking*.
- Rabin, Matthew (2013), "An Approach to Incorporating Psychology into Economics." *American Economic Review*, 103, 617–622.
- Rabin, Matthew and Dimitri Vayanos (2010), "The Gambler's and Hot-Hand Fallacies: Theory and Applications." *Review of Economic Studies*, 77, 730–778.
- Ramey, Valerie A. (2016), "Macroeconomic Shocks and Their Propagation." *Handbook of Macroeconomics*, 2, 71–162.
- Sargent, Thomas J. (1993), *Bounded Rationality in Macroeconomics*. Oxford University Press.
- Sargent, Thomas J. (1999), *The Conquest of American Inflation*. Princeton University Press.
- Sawa, Takamitsu (1978), "Information Criteria for Discriminating Among Alternative Regression Models." *Econometrica*, 46, 1273–1291.
- Shalizi, Cosma Rohilla (2009), "Dynamics of Bayesian Updating with Dependent Data and Misspecified Models." *Electronic Journal of Statistics*, 3, 1039–1074.
- Shimer, Robert (2005), "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95, 25–49.
- Silverman, Leonard M. (1976), "Discrete Riccati Equations: Alternative Algorithms, Asymptotic Properties, and System Theory Interpretations." In *Control and Dynamic Systems*, volume 12, 313–386, Elsevier.
- Sims, Christopher A. (2003), "Implications of Rational Inattention." *Journal of Monetary Economics*, 50, 665–690.
- Stock, James H. and Mark Watson (2011), "Dynamic Factor Models." *Oxford Handbooks Online*.
- Stock, James H. and Mark W. Watson (2002), "Forecasting Using Principal Components From a Large Number of Predictors." *Journal of the American Statistical Association*, 97, 1167–1179.
- Stock, James H. and Mark W. Watson (2016), *Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics*, first edition, volume 2. Elsevier B.V.
- Woodford, Michael (2003), "Imperfect Common Knowledge and the Effects of Monetary Policy." *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, 25.
- Woodford, Michael (2013), "Macroeconomic Analysis Without the Rational Expectations Hypothesis." *Annual Review of Economics*, 5, 303–346.

Online Appendices

C Proofs

Proof of Theorem 2

As a preliminary step, I fix an arbitrary d -state model $\theta = (A, B, Q, R)$ for the agent and compute her forecasts and the KLDR of her model from the true process. If the support of P^θ does not coincide with \mathcal{W} , the support of the true process, then $\text{KLDR}(\theta) = +\infty$. In what follows, I assume that P^θ is supported on \mathcal{W} .

The Kullback–Leibler divergence rate. Since the entropy rate of the true process is finite, the KLDR of θ from the true process is given by

$$\text{KLDR}(\theta) = \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} [-\log f^\theta(y_1, \dots, y_t)] + \text{constant}.$$

Furthermore, by the chain rule,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} [-\log f^\theta(y_1, \dots, y_t)] = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \mathbb{E} [-\log f^\theta(y_\tau | y_{\tau-1}, \dots, y_1)].$$

Since P^θ and \mathbb{P} are both stationary,

$$\mathbb{E} [-\log f^\theta(y_\tau | y_{\tau-1}, \dots, y_1)] = \mathbb{E} [-\log f^\theta(y_0 | y_{-1}, \dots, y_{1-\tau})].$$

On the other hand, since P^θ is a stationary ergodic Gaussian process and $\mathbb{E} [\|y_t\|^2] < \infty$, the sequence $\{-\log f^\theta(y_0 | y_{-1}, \dots, y_{1-\tau})\}_\tau$ is uniformly bounded by an integrable function for any θ . Thus, by the dominated convergence theorem,

$$\lim_{\tau \rightarrow \infty} \mathbb{E} [-\log f^\theta(y_\tau | y_{\tau-1}, \dots, y_1)] = \mathbb{E} [-\log f^\theta(y_0 | y_{-1}, \dots)] = \mathbb{E} [-\log f^\theta(y_{t+1} | y_t, \dots)],$$

where the second equality uses the stationarity of P^θ . The above display implies that the Cesàro sum also converges:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \mathbb{E} [-\log f^\theta(y_\tau | y_{\tau-1}, \dots, y_1)] \rightarrow \mathbb{E} [-\log f^\theta(y_{t+1} | y_t, \dots)].$$

Therefore, to compute the KLDR, I only need to compute the subjective distribution of y_{t+1} under model θ conditional on the history of observations $\{y_t, y_{t-1}, \dots\}$.

Let $E_t^\theta[\cdot]$ denote the agent's subjective expectation given model θ and conditional on history $\{y_\tau\}_{\tau=-\infty}^t$, and let $\text{Var}_t^\theta(\cdot)$ denote the corresponding variance-covariance matrix. Let $\hat{z}_t \equiv E_t^\theta[z_{t+1}]$ denote the agent's conditional expectation of the subjective state. I can express \hat{z}_t recursively using the Kalman filter:

$$\hat{z}_t = (A - KB')\hat{z}_{t-1} + Ky_t, \tag{C.1}$$

where $K \in \mathbb{R}^{d \times n}$ is the Kalman gain defined as

$$K \equiv A \hat{\Sigma}_z B (B' \hat{\Sigma}_z B + R)^\dagger, \quad (\text{C.2})$$

the dagger denotes the Moore–Penrose pseudo-inverse, and $\hat{\Sigma}_z \equiv \text{Var}_t^\theta(z_{t+1})$ is the subjective conditional variance of z_{t+1} , which solves the following (generalized) algebraic Riccati equation:³⁹

$$\hat{\Sigma}_z = A \left(\hat{\Sigma}_z - \hat{\Sigma}_z B (B' \hat{\Sigma}_z B + R)^\dagger B' \hat{\Sigma}_z \right) A' + Q. \quad (\text{C.3})$$

Solving equation (C.1) backward, I get

$$\hat{z}_t = \sum_{\tau=0}^{\infty} (A - KB')^\tau K y_{t-\tau}.$$

The agent's subjective conditional expectation of y_{t+1} can be written in terms of her conditional expectation of z_{t+1} :

$$E_t^\theta[y_{t+1}] = B' E_t^\theta[z_{t+1}] = B' \sum_{\tau=0}^{\infty} (A - KB')^\tau K y_{t-\tau}.$$

Likewise, the conditional variance of y_{t+1} can be expressed in terms of the conditional variance of z_{t+1} :

$$\hat{\Sigma}_y \equiv \text{Var}_t^\theta(y_{t+1}) = B' \hat{\Sigma}_z B + R. \quad (\text{C.4})$$

More generally, the agent's s -period ahead forecast of the vector of observables is given by

$$E_t^\theta[y_{t+s}] = B' A^{s-1} E_t^\theta[z_{t+1}] = B' A^{s-1} \sum_{\tau=0}^{\infty} (A - KB')^\tau K y_{t-\tau}. \quad (\text{C.5})$$

The Kullback–Leibler divergence rate is thus equal to

$$\begin{aligned} \text{KLDR}(\theta) = & -\frac{1}{2} \log \det^* (\hat{\Sigma}_y^\dagger) + \frac{n}{2} \log(2\pi) + \frac{1}{2} \text{tr} (\hat{\Sigma}_y^\dagger \Gamma_0) \\ & - \frac{1}{2} \sum_{\tau=1}^{\infty} \text{tr} (\hat{\Sigma}_y^\dagger \Phi_\tau \Gamma_\tau') - \frac{1}{2} \sum_{\tau=1}^{\infty} \text{tr} (\hat{\Sigma}_y^\dagger \Gamma_\tau \Phi_\tau') \\ & + \frac{1}{2} \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} \text{tr} (\hat{\Sigma}_y^\dagger \Phi_s \Gamma_{\tau-s} \Phi_\tau') + \text{constant}, \end{aligned} \quad (\text{C.6})$$

where $\Gamma_l \equiv \mathbb{E}[y_t y_{t-l}']$ denotes the lag- l autocovariance matrix for the vector of observables under the true process, $\Phi_\tau \equiv B'(A - KB')^{\tau-1} K$, and the constant contains terms that do not depend on θ . Matrix $\hat{\Sigma}_y^\dagger$ denotes the Moore–Penrose pseudo-inverse of $\hat{\Sigma}_y$ and $\det^*(\hat{\Sigma}_y^\dagger)$ denotes its pseudo-determinant.⁴⁰ These objects are the appropriate counterparts of the matrix inverse and the determinant for the case where \mathcal{W} does not equal \mathbb{R}^n , and so, the subjective model θ is degenerate.

³⁹Note that I allow for the possibility that P^θ is supported on some proper subspace \mathcal{W} of \mathbb{R}^n , in which case $B' \hat{\Sigma}_z B + R$ might not be invertible. The Moore–Penrose pseudo-inverse is then the appropriate generalization of matrix inverse in the expression for the Kalman gain. See Chapter 4 of [Anderson and Moore \(2005\)](#) for a treatment in the non-singular case and [Silverman \(1976\)](#) for the case where $B' \hat{\Sigma}_z B + R$ may be singular.

⁴⁰The pseudo-determinant is the product of all non-zero eigenvalues of a square matrix.

Proof of Theorem 2. Let \tilde{n} denote the dimension of vector $\tilde{y}_t = Ty_t$, let $\widetilde{\mathcal{W}}$ denote the linear subspace of $\mathbb{R}^{\tilde{n}}$ defined as $\widetilde{\mathcal{W}} \equiv \{\tilde{y} \in \mathbb{R}^{\tilde{n}} : \tilde{y} = Ty \text{ for some } y \in \mathcal{W}\}$, let $\widetilde{\Theta}_d$ denote the set of d -state models when the vector of observable is $\tilde{y}_t \in \mathbb{R}^{\tilde{n}}$, let $\widetilde{\text{KLDR}}(\tilde{\theta})$ denote the KLDR of model $\tilde{\theta} \in \widetilde{\Theta}_d$ from the true process $\tilde{\mathbb{P}} \equiv \mathbb{P} \circ T^{-1}$, and let $\widetilde{\Theta}_d^*$ denote the set of models $\tilde{\theta}^* \in \widetilde{\Theta}_d$ such that $\widetilde{\text{KLDR}}(\tilde{\theta}^*) \leq \widetilde{\text{KLDR}}(\tilde{\theta})$ for all $\tilde{\theta} \in \widetilde{\Theta}_d$.

I first show that $\widetilde{\mathcal{W}}$ is both the support of any distribution in the set $\mathcal{P}_d^* \circ T^{-1}$ of distributions over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$ induced by \mathcal{P}_d^* and T and the support of any distribution in set $\widetilde{\mathcal{P}}_d^*$. Note that there always exists a d -state model θ for which $\text{KLDR}(\theta) < \infty$ —one such model is the one according to which y_t is i.i.d. over time and has a variance-covariance matrix that coincides with the true variance-covariance matrix Γ_0 . Therefore, for any pseudo-true d -state model, the KLDR is finite. Thus, any process $P \in \mathcal{P}_d^*$ is supported on \mathcal{W} , and so, any process $P \in \mathcal{P}_d^* \circ T^{-1}$ is supported on $\widetilde{\mathcal{W}}$. On the other hand, since the true distribution \mathbb{P} is supported on \mathcal{W} , the induced distribution $\tilde{\mathbb{P}} \equiv \mathbb{P} \circ T^{-1}$ is supported on $\widetilde{\mathcal{W}}$. Consequently, by the above argument, any distribution $P \in \widetilde{\mathcal{P}}_d^*$ is also supported on $\widetilde{\mathcal{W}}$. Therefore, I can restrict my attention to models $\theta \in \Theta_d$ such that P^θ is supported on \mathcal{W} and models $\tilde{\theta} \in \widetilde{\Theta}_d$ such that $P^{\tilde{\theta}}$ is supported on $\widetilde{\mathcal{W}}$.

For any model $\theta = (A, B, Q, R) \in \Theta_d$, define model $T(\theta) \in \widetilde{\Theta}_d$ as $T(\theta) \equiv (A, BT', Q, TRT')$. I next show that $\widetilde{\text{KLDR}}(T(\theta)) = \text{KLDR}(\theta)$, up to an additive constant that does not depend on θ . Fix some model $\theta \in \Theta_d$. Let $\hat{\Sigma}_z \equiv \text{Var}_t^\theta(z_{t+1})$ denote the subjective conditional variance of the subjective state under model θ , and let $\widetilde{\Sigma}_z \equiv \text{Var}_t^{T(\theta)}(z_{t+1})$ denote the corresponding conditional variance under model $T(\theta)$. Matrices $\hat{\Sigma}_z$ and $\widetilde{\Sigma}_z$ solve the following Riccati equations:

$$\hat{\Sigma}_z = A \left(\hat{\Sigma}_z - \hat{\Sigma}_z B (B' \hat{\Sigma}_z B + R)^\dagger B' \hat{\Sigma}_z \right) A' + Q, \quad (\text{C.7})$$

$$\widetilde{\Sigma}_z = A \left(\widetilde{\Sigma}_z - \widetilde{\Sigma}_z BT' \left(TB' \widetilde{\Sigma}_z BT' + TRT' \right)^\dagger TB' \widetilde{\Sigma}_z \right) A' + Q. \quad (\text{C.8})$$

Since matrix T has full rank, $T^\dagger = (T'T)^{-1}T$ and $T^\dagger T = I$. Therefore, $\widetilde{\Sigma}_z = \hat{\Sigma}_z$. Next, let K denote the Kalman gain given model θ , and let \tilde{K} denote the Kalman gain given model $T(\theta)$. Note that

$$\tilde{K} = A \widetilde{\Sigma}_z BT' \left(TB' \widetilde{\Sigma}_z BT' + TRT' \right)^\dagger = KT^\dagger.$$

Let $\Phi_\tau \equiv B'(A - KB')^{\tau-1}K$, and let $\widetilde{\Phi}_\tau$ denote the corresponding matrix given model $T(\theta)$. Note that

$$\widetilde{\Phi}_\tau \equiv TB'(A - KT^\dagger TB')^{\tau-1}KT^\dagger = T\Phi_\tau T^\dagger.$$

Finally, let $\hat{\Sigma}_y \equiv \text{Var}_t^\theta(y_{t+1})$ denote the subjective conditional variance of y_{t+1} given model θ , and let $\widetilde{\Sigma}_y \equiv \text{Var}_t^{T(\theta)}(\tilde{y}_{t+1})$ denote the corresponding conditional variance given model $T(\theta)$. Note that

$$\widetilde{\Sigma}_y = TB' \widetilde{\Sigma}_z BT' + TRT' = T\hat{\Sigma}_y T'.$$

One the other hand, $\tilde{\Gamma}_l \equiv \tilde{\mathbb{E}}[\tilde{y}_t \tilde{y}_{t-l}'] = T\mathbb{E}[y_t y_{t-l}']T' = T\Gamma_l T'$. Therefore, by equation (C.6),

$$\widetilde{\text{KLDR}}(T(\theta)) = -\frac{1}{2} \log \det^* \left(T^{\dagger'} \hat{\Sigma}_y^\dagger T^\dagger \right) + \frac{n}{2} \log(2\pi) + \frac{1}{2} \text{tr} \left(T^{\dagger'} \hat{\Sigma}_y^\dagger T^\dagger T \Gamma_0 T' \right)$$

$$\begin{aligned}
& -\frac{1}{2} \sum_{\tau=1}^{\infty} \text{tr} \left(T^{\dagger'} \hat{\Sigma}_y^{\dagger} T^{\dagger} T \Phi_{\tau} T^{\dagger} T \Gamma_{\tau}' T' \right) - \frac{1}{2} \sum_{\tau=1}^{\infty} \text{tr} \left(T^{\dagger'} \hat{\Sigma}_y^{\dagger} T^{\dagger} T \Gamma_{\tau} T' T^{\dagger'} \Phi_{\tau}' T' \right) \\
& + \frac{1}{2} \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} \text{tr} \left(T^{\dagger'} \hat{\Sigma}_y^{\dagger} T^{\dagger} T \Phi_s T^{\dagger} T \Gamma_{\tau-s} T' T^{\dagger'} \Phi_{\tau}' T' \right) + \text{constant}.
\end{aligned}$$

The fact that $T^{\dagger} T = I$ implies that the above expression is equal to $\text{KLDR}(\theta)$, up to an additive constant that does not depend on θ .

Likewise, for any model $\tilde{\theta} = (\tilde{A}, \tilde{B}, \tilde{Q}, \tilde{R}) \in \tilde{\Theta}_d$, define $T^{-1}(\tilde{\theta}) \equiv (\tilde{A}, \tilde{B} T^{\dagger'}, \tilde{Q}, T^{\dagger} \tilde{R} T^{\dagger'}) \in \Theta_d$. By an argument similar to the one in the previous paragraph, $\text{KLDR}(T^{-1}(\tilde{\theta})) = \widetilde{\text{KLDR}}(\tilde{\theta})$, up to an additive constant that does not depend on $\tilde{\theta}$.

Therefore, the mapping T defines an isomorphism between the set of models Θ_d and the set of models $\tilde{\Theta}_d$: Any model $\theta \in \Theta_d$ can be identified with a model $T(\theta) \in \tilde{\Theta}_d$ such that the KLDR of P^{θ} from the process \mathbb{P} is equal to the KLDR of $P^{T(\theta)}$ from $\mathbb{P} \circ T^{-1}$, and any model $\tilde{\theta} \in \tilde{\Theta}_d$ can be identified with a model $T^{-1}(\tilde{\theta}) \in \Theta_d$ such that the KLDR of $P^{T^{-1}(\tilde{\theta})}$ from the process \mathbb{P} is equal to the KLDR of $P^{\tilde{\theta}}$ from the process $\mathbb{P} \circ T^{-1}$. This conclusion immediately implies that the set of pseudo-true d -state models under true process \mathbb{P} is identified with the set of pseudo-true d -state models under true process $\mathbb{P} \circ T^{-1}$. That is, $\tilde{\Theta}_d^* = \{T(\theta) : \theta \in \Theta_d^*\}$.

It only remains to show that $P^{T(\theta)} = P^{\theta} \circ T^{-1}$ for any model $\theta \in \Theta_d$. Since $P^{T(\theta)}$ and $P^{\theta} \circ T^{-1}$ are both zero mean, stationary, and normal distributions over $\{\tilde{y}_t\}_{t=-\infty}^{\infty}$, it is sufficient to show that the autocovariance matrices of \tilde{y}_t are identical at all lags under the two distributions. But this follows the definitions of distributions $P^{T(\theta)}$ and $P^{\theta} \circ T^{-1}$. \square

Proof of Theorem 3

Before establishing the theorem, I state and prove a lemma that underpins all the characterization results of the paper:

Lemma C.1. *Model $\theta = (A, B, Q, R)$ is a pseudo-true d -state model given true autocovariance matrices $\{\Gamma_l\}_l$ with Γ_0 invertible if and only if $A = M$, $B = D' N^{-1}$, $Q = I - M(I - D'D)M'$, and $R = N^{-1'}(I - DD')N^{-1}$, where (M, D, N) is a tuple that minimizes*

$$\begin{aligned}
\text{KLDR}(\tilde{M}, \tilde{D}, \tilde{N}) \equiv & -\frac{1}{2} \log \det(\tilde{N} \tilde{N}') + \frac{1}{2} \text{tr}(\tilde{N}' \Gamma_0 \tilde{N}) - \sum_{\tau=1}^{\infty} \text{tr} \left((\tilde{M}(I - \tilde{D}' \tilde{D}))^{\tau-1} \tilde{M} \tilde{D}' \tilde{N}' \Gamma_{\tau}' \tilde{N} \tilde{D} \right) \\
& + \frac{1}{2} \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} \text{tr} \left(\tilde{D} (\tilde{M}(I - \tilde{D}' \tilde{D}))^{s-1} \tilde{M} \tilde{D}' \tilde{N}' \Gamma_{\tau-s} \tilde{N} \tilde{D} \tilde{M}' ((I - \tilde{D}' \tilde{D}) \tilde{M}')^{\tau-1} \tilde{D}' \right)
\end{aligned} \tag{C.9}$$

subject to the constraints that \tilde{M} is a $d \times d$ convergent matrix, \tilde{D} is an $n \times d$ diagonal matrix with elements in the $[0, 1]$ interval, \tilde{N} is an $n \times n$ invertible matrix, and $\|\tilde{M}(I - \tilde{D}' \tilde{D}) \tilde{M}'\|_2 < 1$.

Proof. The assumption that Q is positive definite implies that the solution $\hat{\Sigma}_z$ to the Riccati equation (C.3) is invertible. On the other hand, since Γ_0 is invertible, I can restrict attention to sub-

jective models for which $\hat{\Sigma}_y$ is non-singular.⁴¹ The pseudo-inverses and pseudo-determinants in equations (C.3) and (C.6) thus reduce to matrix inverses and determinants.

I start by expressing $\hat{\Sigma}_y^{-\frac{1}{2}} B' \hat{\Sigma}_z^{\frac{1}{2}}$ as its singular value decomposition:

$$\hat{\Sigma}_y^{-\frac{1}{2}} B' \hat{\Sigma}_z^{\frac{1}{2}} = U D V', \quad (\text{C.10})$$

where $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{d \times d}$ are orthogonal matrices, and $D \in \mathbb{R}^{n \times d}$ is a rectangular diagonal matrix with singular values of $\hat{\Sigma}_z^{\frac{1}{2}} B' \hat{\Sigma}_y^{-\frac{1}{2}}$ on the diagonal. Note that

$$V D' D V' = \hat{\Sigma}_z^{\frac{1}{2}} B (B' \hat{\Sigma}_z B + R)^{-1} B' \hat{\Sigma}_z^{\frac{1}{2}}.$$

Since R is a symmetric positive semidefinite matrix and V is orthogonal, diagonal elements of D are weakly smaller than 1 (strictly so if R is positive definite). Next define

$$M \equiv V^{-1} \hat{\Sigma}_z^{-\frac{1}{2}} A \hat{\Sigma}_z^{\frac{1}{2}} V.$$

Then,

$$A = \hat{\Sigma}_z^{\frac{1}{2}} V M V^{-1} \hat{\Sigma}_z^{-\frac{1}{2}}, \quad (\text{C.11})$$

$$B = \hat{\Sigma}_z^{-\frac{1}{2}} V D' U' \hat{\Sigma}_y^{\frac{1}{2}}, \quad (\text{C.12})$$

$$K = \hat{\Sigma}_z^{\frac{1}{2}} V M D' U' \hat{\Sigma}_y^{-\frac{1}{2}}, \quad (\text{C.13})$$

and so

$$\begin{aligned} K B' &= \hat{\Sigma}_z^{\frac{1}{2}} V M D' D V' \hat{\Sigma}_z^{-\frac{1}{2}}, \\ \Phi_\tau &= \hat{\Sigma}_y^{\frac{1}{2}} U D (M (I - D' D))^{\tau-1} M D' U' \hat{\Sigma}_y^{-\frac{1}{2}}. \end{aligned}$$

Note that since A is a convergent matrix, so is M . Substituting in (C.3) for A from equation (C.11) and for B from (C.12), I get

$$\begin{aligned} Q &= \hat{\Sigma}_z - A \left(\hat{\Sigma}_z - \hat{\Sigma}_z B (B' \hat{\Sigma}_z B + R)^{-1} B' \hat{\Sigma}_z \right) A' \\ &= \hat{\Sigma}_z - \hat{\Sigma}_z^{\frac{1}{2}} V M V^{-1} \hat{\Sigma}_z^{-\frac{1}{2}} \left(\hat{\Sigma}_z - \hat{\Sigma}_z^{\frac{1}{2}} V D' D V \hat{\Sigma}_z^{\frac{1}{2}} \right) \hat{\Sigma}_z^{-\frac{1}{2}} V M' V^{-1} \hat{\Sigma}_z^{\frac{1}{2}} \\ &= \hat{\Sigma}_z - \hat{\Sigma}_z^{\frac{1}{2}} V M (I - D' D) M' V^{-1} \hat{\Sigma}_z^{\frac{1}{2}}. \end{aligned} \quad (\text{C.14})$$

Therefore, since Q is positive definite, the eigenvalues of $V M (I - D' D) M' V^{-1}$ must all lie inside the unit circle. This implies that $\rho(M (I - D' D) M') = \|M (I - D' D) M'\|_2 < 1$, where $\rho(\cdot)$ denotes the spectral radius, and I am using the facts that the spectral radius is invariant to similarity transformations and equal to the spectral norm for symmetric matrices.

I can further reduce the number of parameters in the agent's model by transforming $\hat{\Sigma}_y^{-\frac{1}{2}}$ using the orthogonal matrix U . Define

$$N \equiv \hat{\Sigma}_y^{-\frac{1}{2}} U.$$

⁴¹ Since the variance-covariance matrix Γ_0 of the true process is invertible, $\text{KLDR}(\theta) = +\infty$ for any subjective model θ with a singular $\hat{\Sigma}_y$. Note that, in light of Theorem 2, the restriction to true processes with invertible variance-covariance matrices is without loss of generality.

Since $\hat{\Sigma}_y^{-\frac{1}{2}}$ and U are invertible matrices, so is N . Because $\hat{\Sigma}_y^{-\frac{1}{2}}$ is symmetric,

$$UN' = NU' = \hat{\Sigma}_y^{-\frac{1}{2}},$$

so

$$\hat{\Sigma}_y^{-1} = NU'UN' = NN',$$

and

$$\text{tr} \left(\hat{\Sigma}_y^{-1} \Gamma_0 \right) = \text{tr} \left(\hat{\Sigma}_y^{-\frac{1}{2}} \Gamma_0 \hat{\Sigma}_y^{-\frac{1}{2}} \right) = \text{tr} (UN' \Gamma_0 NU') = \text{tr} (N' \Gamma_0 N).$$

On the other hand,

$$\begin{aligned} \text{tr} \left(\hat{\Sigma}_y^{-1} \Phi_\tau \Gamma'_\tau \right) &= \text{tr} \left(\hat{\Sigma}_y^{-\frac{1}{2}} U D (M (I - D' D))^{\tau-1} M D' U' \hat{\Sigma}_y^{-\frac{1}{2}} \Gamma'_\tau \right) \\ &= \text{tr} \left((M (I - D' D))^{\tau-1} M D' N' \Gamma'_\tau N D \right), \end{aligned}$$

and

$$\text{tr} \left(\hat{\Sigma}_y^{-1} \Phi_s \Gamma_{\tau-s} \Phi'_\tau \right) = \text{tr} \left(D (M (I - D' D))^{s-1} M D' N' \Gamma_{\tau-s} N D M' ((I - D' D) M')^{\tau-1} D' \right).$$

Therefore, the KLDR can be expressed in terms of matrices M , D , and N as

$$\text{KLDR}(\theta) = \text{KLDR}(M, D, N) + \text{constant},$$

where $\text{KLDR}(M, D, N)$ is as in the statement of the lemma.

It only remains to show that, for any $(\hat{M}, \hat{D}, \hat{N})$ such that \hat{M} is a $d \times d$ convergent matrix, \hat{D} is an $n \times d$ diagonal matrix with elements in the $[0, 1]$ interval, \hat{N} is an $n \times n$ invertible matrix, and $\|\hat{M}(I - \hat{D}\hat{D}')\hat{M}'\|_2 < 1$, one can construct a corresponding (A, B, Q, R) such that A is convergent, Q is positive definite, and R is positive semidefinite. Given such a tuple $(\hat{M}, \hat{D}, \hat{N})$, let

$$\begin{aligned} A &= \hat{M}, \\ B &= \hat{D}' \hat{N}^{-1}, \\ Q &= I - \hat{M} (I - \hat{D}' \hat{D}) \hat{M}', \\ R &= \hat{N}^{-1'} (I - \hat{D} \hat{D}') \hat{N}^{-1}. \end{aligned}$$

Since \hat{M} is convergent, so is A . Since $\|\hat{M}(I - \hat{D}\hat{D}')\hat{M}'\|_2 < 1$, matrix Q is positive definite. And since \hat{D} is a diagonal matrix with elements in the $[0, 1]$ interval, R is positive semidefinite. It is easy to verify that then $\hat{\Sigma}_z = I$ is then the solution to the Riccati equation (C.3), and so, $\hat{\Sigma}_y = (\hat{N} \hat{N}')^{-1}$. Therefore, I can choose $U = (\hat{N} \hat{N}')^{-\frac{1}{2}} \hat{N}$, $D = \hat{D}$, and $V = I$ in equation (C.10). Substituting in the expressions for M and N , I get $M = \hat{M}$ and $N = \hat{N}$. This completes the proof of the lemma.

For future reference, I also compute several other objects in terms of the M , D , and N matrices. The matrix of Kalman gain is given by

$$K = M D' N'. \tag{C.15}$$

The subjective forecasts can then be found by substituting for A , B , and K in (C.5):

$$E_t^\theta[y_{t+s}] = N'^{-1}DM^{s-1} \sum_{\tau=0}^{\infty} (M(I - D'D))^\tau MD'N'y_{t-\tau}. \quad (\text{C.16})$$

The subjective variance of y_{t+1} conditional on the information available to the agent at time t is given by

$$\hat{\Sigma}_y = (NN')^{-1}.$$

The unconditional subjective variance of y is given by

$$\text{Var}^\theta(y) = B'\text{Var}^\theta(z)B + R,$$

where $\text{Var}^\theta(z)$ solves the discrete Lyapunov equation

$$\text{Var}^\theta(z) = A\text{Var}^\theta(z)A' + Q.$$

Solving the above equation forward, I get

$$\text{Var}^\theta(z) = I + \sum_{\tau=1}^{\infty} M^\tau D' D M'^\tau.$$

Therefore,

$$\text{Var}^\theta(y) = B' \sum_{\tau=0}^{\infty} A^\tau Q A'^\tau B + R = N^{-1'} \left(I + \sum_{\tau=1}^{\infty} D M^\tau D' D M'^\tau D' \right) N^{-1}. \quad (\text{C.17})$$

□

I can now establish Theorem 3.

Proof of Theorem 3. Let M , D , and N be as in Lemma C.1. When $d = 1$, then

$$M = a$$

for some $a \in [-1, 1]$ and

$$D = \begin{pmatrix} d_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = d_1 e_1$$

for some $d_1 \in [0, 1]$, where e_1 denotes the first coordinate vector. Define

$$\eta \equiv 1 - d_1^2, \\ S \equiv \Gamma_0^{\frac{1}{2}} N.$$

Then KLDR, defined in (C.9), can be written (with slight abuse of notation) as a function of a , η , and S :

$$\text{KLDR}(a, \eta, S) = -\frac{1}{2} \log \det(SS') + \frac{1}{2} \text{tr}(S'S) - \frac{1}{2} e_1' S' \Omega(a, \eta) S e_1,$$

where

$$\Omega(a, \eta) \equiv a(1 - \eta) \sum_{\tau=1}^{\infty} (a\eta)^{\tau-1} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau} + \Gamma'_{\tau}) \Gamma_0^{-\frac{1}{2}} - a^2(1 - \eta)^2 \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} (a\eta)^{s+\tau-2} \Gamma_0^{-\frac{1}{2}} \Gamma_{\tau-s} \Gamma_0^{-\frac{1}{2}}.$$

I can simplify the second term of $\Omega(a, \eta)$ further:

$$\begin{aligned} \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} (a\eta)^{s+\tau-2} \Gamma_0^{-\frac{1}{2}} \Gamma_{\tau-s} \Gamma_0^{-\frac{1}{2}} &= \sum_{s=1}^{\infty} \sum_{\tau=s+1}^{\infty} (a\eta)^{s+\tau-2} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau-s} + \Gamma'_{\tau-s}) \Gamma_0^{-\frac{1}{2}} + \sum_{s=1}^{\infty} (a\eta)^{2(s-1)} I \\ &= \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} (a\eta)^{2(s-1)+\tau} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau} + \Gamma'_{\tau}) \Gamma_0^{-\frac{1}{2}} + \sum_{s=1}^{\infty} (a\eta)^{2(s-1)} I \\ &= \left(\sum_{s=1}^{\infty} (a\eta)^{2(s-1)} \right) \left(I + \sum_{\tau=1}^{\infty} (a\eta)^{\tau} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau} + \Gamma'_{\tau}) \Gamma_0^{-\frac{1}{2}} \right) \\ &= \frac{1}{1 - a^2 \eta^2} \left(I + a\eta \sum_{\tau=1}^{\infty} (a\eta)^{\tau-1} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau} + \Gamma'_{\tau}) \Gamma_0^{-\frac{1}{2}} \right). \end{aligned}$$

Therefore,

$$\Omega(a, \eta) = -\frac{a^2(1 - \eta)^2}{1 - a^2 \eta^2} I + \frac{(1 - \eta)(1 - a^2 \eta)}{1 - a^2 \eta^2} \sum_{\tau=1}^{\infty} a^{\tau} \eta^{\tau-1} \Gamma_0^{-\frac{1}{2}} (\Gamma_{\tau} + \Gamma'_{\tau}) \Gamma_0^{-\frac{1}{2}}. \quad (\text{C.18})$$

By Lemma C.1, minimizing the KLDR with respect to A , B , Q , and R is equivalent to minimizing $\text{KLDR}(M, D, N)$ with respect to M , D , and N . But for any a , η , and S , one can construct a corresponding M , D , and N , and vice versa. Therefore, I can instead minimize $\text{KLDR}(a, \eta, S)$ with respect to a , η , and S .

I first minimize $\text{KLDR}(a, \eta, S)$ with respect to S taking a and η as given. The first-order optimality condition with respect to S is given by

$$S^{-1} = S' - e_1 e_1' S' \Omega(a, \eta),$$

which implies that

$$S' S - e_1 e_1' S' \Omega(a, \eta) S = I. \quad (\text{C.19})$$

Therefore, for any solution to the problem of minimizing $\text{KLDR}(a, \eta, S)$,

$$n = \text{tr}(I) = \text{tr}(S' S) - \text{tr}(e_1 e_1' S' \Omega(a, \eta) S) = \text{tr}(S' S) - e_1' S' \Omega(a, \eta) S e_1.$$

Thus, minimizing $\text{KLDR}(a, \eta, S)$ with respect to a , η , and S is equivalent to solving the following program:

$$\max_{a, \eta} \det(S(a, \eta) S'(a, \eta)),$$

where

$$S(a, \eta) \in \arg \min_S -\frac{1}{2} \log \det(SS') + \frac{1}{2} \text{tr}(S' S) - \frac{1}{2} e_1' S' \Omega(a, \eta) S e_1. \quad (\text{C.20})$$

I proceed by first characterizing $S(a, \eta)$. Note that the necessary first-order optimality conditions for problem (C.20) are given by matrix equation (C.19).

Claim C.1. For any matrix S that solves equation (C.19), the necessary first-order optimality condition for problem (C.20),

$$(i) \quad S e_1 = \frac{1}{\sqrt{1-\lambda}} u,$$

$$(ii) \quad S'^{-1} e_1 = \sqrt{1-\lambda} u,$$

$$(iii) \quad S S' = I + \frac{\lambda}{1-\lambda} u u',$$

where λ is an eigenvalue of the real symmetric matrix $\Omega(a, \eta)$ and u is a corresponding eigenvector normalized such that $u'u = 1$.

I return to proving the claim toward the end of the proof. Equation (C.19) in general has multiple solutions, with each solution corresponding to a local extremum of problem (C.20). The global optimum of problem (C.20) is given by the solution to equation (C.19) that results in the largest value for $\det(SS')$. But by part (iii) of Claim C.1, $\det(SS') = (1-\lambda)^{-1}$. Thus, for any pseudo-true one-state model, a and η maximize $\lambda_{\max}(\Omega(a, \eta))$ and S satisfies parts (i)–(iii) of Claim C.1, with $\lambda = \lambda_{\max}(\Omega)$ and $u = u_{\max}(\Omega)$ the corresponding eigenvector.

I next find parameters A, B, Q , and R representing the a, η , and S that minimize $\text{KLDR}(a, \eta, S)$. First, note that

$$\begin{aligned} M &= a, \\ D &= \sqrt{1-\eta} e_1, \\ N &= \Gamma_0^{-\frac{1}{2}} S. \end{aligned}$$

The representation in Lemma C.1 is thus given by⁴²

$$\begin{aligned} A &= a, \\ B &= \sqrt{1-\eta} e_1' S^{-1} \Gamma_0^{\frac{1}{2}}, \\ Q &= 1 - a^2 \eta, \\ R &= \Gamma_0^{\frac{1}{2}} S^{-1'} (I - (1-\eta) e_1 e_1') S^{-1} \Gamma_0^{\frac{1}{2}}. \end{aligned}$$

By Claim C.1 and the argument above,

$$\begin{aligned} e_1' S^{-1} &= \sqrt{1-\lambda_{\max}(\Omega)} u'_{\max}(\Omega), \\ S^{-1'} S^{-1} &= (SS')^{-1} = I - \lambda_{\max}(\Omega) u_{\max}(\Omega) u'_{\max}(\Omega). \end{aligned}$$

Thus,

$$B = \sqrt{(1-\eta)(1-\lambda_{\max}(\Omega))} u'_{\max}(\Omega) \Gamma_0^{\frac{1}{2}},$$

⁴²For this (A, B, Q, R) tuple to represent a one-state model, I need A to be convergent, Q to be positive definite, and R to be positive semidefinite. That R is always positive semidefinite is immediate. Showing that A is convergent and Q is positive definite takes more work. I do so in Theorem 4.

and

$$\begin{aligned} R &= \Gamma_0^{\frac{1}{2}} (I - \lambda_{\max}(\Omega) u_{\max}(\Omega) u'_{\max}(\Omega)) \Gamma_0^{\frac{1}{2}} - (1 - \eta) (I - \lambda_{\max}(\Omega)) \Gamma_0^{\frac{1}{2}} u_{\max}(\Omega) u'_{\max}(\Omega) \Gamma_0^{\frac{1}{2}} \\ &= \Gamma_0^{\frac{1}{2}} [I - (1 - \eta + \eta \lambda_{\max}(\Omega)) u_{\max}(\Omega) u'_{\max}(\Omega)] \Gamma_0^{\frac{1}{2}}. \end{aligned}$$

Finally, note that $M = a$, $D = \sqrt{1 - \eta} e_1$, and $N = \Gamma_0^{-\frac{1}{2}} S$. Therefore, by equation (C.16), the subjective forecasts are given by

$$E_t^\theta [y_{t+s}] = a^s (1 - \eta) \Gamma_0^{\frac{1}{2}} S'^{-1} e_1 e_1' S' \Gamma_0^{-\frac{1}{2}} \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}. \quad (\text{C.21})$$

Using Claim C.1 to substitute for the optimal S , I get

$$E_t^\theta [y_{t+s}] = a^s (1 - \eta) \Gamma_0^{\frac{1}{2}} u_{\max}(\Omega) u'_{\max}(\Omega) \Gamma_0^{-\frac{1}{2}} \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau},$$

where $u_{\max}(\Omega)$ is a unit-norm eigenvector of Ω with eigenvalue $\lambda_{\max}(\Omega)$. The theorem then follows by the definitions of p and q . \square

Proof of Claim C.1. The first-order optimality condition with respect to S is given by

$$S'S - e_1 e_1' S' \Omega S = I. \quad (\text{C.22})$$

Multiplying the transpose of the above equation from right by e_1 and from left by S'^{-1} , I get

$$S e_1 - \Omega S e_1 = S'^{-1} e_1. \quad (\text{C.23})$$

On the other hand, multiplying equation (C.22) from left by S and from right by S^{-1} , I get

$$S S' = I + S e_1 e_1' S' \Omega. \quad (\text{C.24})$$

By the Sherman–Morrison formula,

$$S'^{-1} S^{-1} = I - \frac{S e_1 e_1' S' \Omega}{1 + e_1' S' \Omega S e_1}.$$

Multiplying the above equation from right by $S e_1$, I get

$$S'^{-1} e_1 = \frac{1}{1 + e_1' S' \Omega S e_1} S e_1. \quad (\text{C.25})$$

Substituting for $S'^{-1} e_1$ from the above equation in (C.23) and rearranging the terms, I get

$$\Omega S e_1 = \frac{e_1' S' \Omega S e_1}{1 + e_1' S' \Omega S e_1} S e_1. \quad (\text{C.26})$$

That is, $S e_1$ is an eigenvector of Ω . Let λ denote the corresponding eigenvalue and let $u = S e_1 / \sqrt{e_1' S' S e_1}$. Then equation (C.26) implies

$$\lambda = \frac{\lambda e_1' S' S e_1}{1 + \lambda e_1' S' S e_1}.$$

I separately consider the cases $\lambda \neq 0$ and $\lambda = 0$. If $\lambda \neq 0$, then

$$e_1' S' S e_1 = (1 - \lambda)^{-1},$$

and so

$$S e_1 = \frac{1}{\sqrt{1 - \lambda}} u.$$

Equation (C.25) then implies that

$$S'^{-1} e_1 = \sqrt{1 - \lambda} u,$$

and equation (C.24) implies that

$$S S' = I + \frac{\lambda}{1 - \lambda} u u'.$$

If $\lambda = 0$, then equation (C.23) implies that $S e_1 = S'^{-1} e_1$, and so, $S e_1$ and $S'^{-1} e_1$ are both multiples of u . Furthermore, $e_1' S^{-1} S e_1 = e_1' e_1 = 1$. Therefore, $S e_1 = S'^{-1} e_1 = u$. On the other hand, equation (C.24) implies that $S S' = I$. This completes the proof of the claim. \square

Proof of Theorem 4

I first prove a useful lemma about the spectral radius of autocorrelation matrices:

Lemma C.2. *For any purely non-deterministic, stationary ergodic, and non-degenerate process with autocorrelation matrices $\{C_l\}_l$, the spectral radii of autocorrelation matrices satisfy $\rho(C_l) \leq 1$ for any l with the inequality strict for $l = 1$.*

Proof. Let λ_l denote an eigenvalue of C_l largest in magnitude and let u_l denote the corresponding eigenvector normalized such that $u_l' u_l = 1$. Define the process $\omega_t^{(l)} \equiv u_l' \Gamma_0^{-\frac{1}{2}} y_t \in \mathbb{R}$. Since y_t is a purely non-deterministic, stationary ergodic, and non-degenerate process, so is $\omega_t^{(l)}$ for any l . I first show that λ_l is the autocorrelation of process $\omega_t^{(l)}$ at lag l . Note that

$$\mathbb{E}[\omega_t^{(l)} \omega_{t-l}^{(l)}] = u_l' \Gamma_0^{-\frac{1}{2}} \mathbb{E}[y_t y_{t-l}'] \Gamma_0^{-\frac{1}{2}} u_l = u_l' \Gamma_0^{-\frac{1}{2}} \Gamma_l \Gamma_0^{-\frac{1}{2}} u_l = u_l' \Gamma_0^{-\frac{1}{2}} \left(\frac{\Gamma_l + \Gamma_l'}{2} \right) \Gamma_0^{-\frac{1}{2}} u_l = u_l' C_l u_l = \lambda_l.$$

Furthermore,

$$\mathbb{E}[\omega_t^{(l)} \omega_t^{(l)}] = u_l' \Gamma_0^{-\frac{1}{2}} \mathbb{E}[y_t y_t'] \Gamma_0^{-\frac{1}{2}} u_l = u_l' \Gamma_0^{-\frac{1}{2}} \Gamma_0 \Gamma_0^{-\frac{1}{2}} u_l = u_l' u_l = 1.$$

Therefore, since $\omega_t^{(l)}$ is purely non-deterministic, stationary ergodic, and non-degenerate,

$$\rho(C_l) = |\lambda_l| = \frac{\mathbb{E}[\omega_t^{(l)} \omega_{t-l}^{(l)}]}{\mathbb{E}[\omega_t^{(l)} \omega_t^{(l)}]} \leq 1.$$

Next, toward a contradiction suppose that $\rho(C_1) = 1$. Then $\omega_t^{(1)}$ is perfectly correlated with $\omega_{t-1}^{(1)}$, and so, with $\omega_{t-l}^{(1)}$ for every l , contradicting the assumption that $\omega_t^{(1)}$ is purely non-deterministic, stationary ergodic, and non-degenerate. \square

I can now prove the theorem.

Proof of Theorem 4. Define

$$C(a, \eta) \equiv \sum_{\tau=1}^{\infty} a^{\tau} \eta^{\tau-1} C_{\tau}. \quad (\text{C.27})$$

Then

$$\lambda_{\max}(\Omega(a, \eta)) = -\frac{a^2(1-\eta)^2}{1-a^2\eta^2} + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2} \lambda_{\max}(C(a, \eta)), \quad (\text{C.28})$$

where $\lambda_{\max}(C(a, \eta))$ denotes the largest eigenvalue of $C(a, \eta)$. To simplify the exposition, I prove the result under the assumption that the largest eigenvalue of $C(a, \eta)$ is simple at the point (a^*, η^*) that maximizes $\lambda_{\max}(\Omega(a, \eta))$.⁴³ The partial derivatives of $\lambda_{\max}(\Omega(a, \eta))$ with respect to a and η are given by

$$\begin{aligned} \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a} &= \frac{-2a(1-\eta)^2}{(1-a^2\eta^2)^2} - \frac{4a\eta(1-\eta)^2}{(1-a^2\eta^2)^2} \lambda_{\max}(C) \\ &\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2} u'_{\max}(C) \frac{\partial C}{\partial a} u_{\max}(C), \end{aligned} \quad (\text{C.29})$$

$$\begin{aligned} \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta} &= \frac{2a^2(1-\eta)(1-a^2\eta)}{(1-a^2\eta^2)^2} - \frac{2(1+a^4\eta^2+a^2(1-4\eta+\eta^2))}{(1-a^2\eta^2)^2} \lambda_{\max}(C) \\ &\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2} u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C), \end{aligned} \quad (\text{C.30})$$

where $u_{\max}(C)$ denotes the eigenvector of C with eigenvalue $\lambda_{\max}(C)$, normalized such that $u'_{\max}(C)u_{\max}(C) = 1$, and

$$\begin{aligned} \frac{\partial C}{\partial a} &= \sum_{\tau=1}^{\infty} \tau a^{\tau-1} \eta^{\tau-1} C_{\tau}, \\ \frac{\partial C}{\partial \eta} &= \sum_{\tau=1}^{\infty} (\tau-1) a^{\tau} \eta^{\tau-2} C_{\tau}. \end{aligned}$$

Note that

$$\eta u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C) + \lambda_{\max}(C) = a u'_{\max}(C) \frac{\partial C}{\partial a} u_{\max}(C). \quad (\text{C.31})$$

for any a and η .

Let a^* and η^* be scalars in the $[-1, 1]$ and $[0, 1]$ intervals, respectively, that maximize $\lambda_{\max}(\Omega(a, \eta))$. I separately consider the cases $\eta^* = 1$ and $\eta^* < 1$. If $\eta^* = 1$, then $B = 0$ in the representation in the proof of Theorem 3, the pseudo-true one-state model is i.i.d., and $A = a^*$ can be chosen arbitrarily to satisfy $|a^*| < 1$.⁴⁴

In the rest of the proof, I assume that $\eta^* < 1$ and show that this implies $a^* \neq 1$ —by a similar argument $a^* \neq -1$. Toward a contradiction, suppose $a^* = 1$. Setting $a = 1$ in the partial derivatives

⁴³The argument can easily be adapted to the case where the largest eigenvalue of $C(a^*, \eta^*)$ is not necessarily simple by replacing the gradient of $\lambda_{\max}(C(a, \eta))$ with its subdifferential and replacing the usual first-order optimality condition with the condition that the zero vector belongs to the subdifferential.

⁴⁴The pseudo-true one-state model then has a zero-state minimal representation.

of $\lambda_{\max}(\Omega(a, \eta))$, I get

$$\begin{aligned}\left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a} \right|_{a=1} &= \frac{2(1-\eta)^2}{(1-\eta^2)^2} \left[-1 - 2\eta \lambda_{\max}(C) + (1-\eta^2) u'_{\max}(C) \frac{\partial C}{\partial a} u_{\max}(C) \right], \\ \left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta} \right|_{a=1} &= \frac{2(1-\eta)^2}{(1-\eta^2)^2} \left[1 - 2\lambda_{\max}(C) + (1-\eta^2) u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C) \right],\end{aligned}$$

where $C = C(1, \eta)$ and its partial derivatives are computed at $a = 1$. Multiplying the second equation above by η and subtracting from it the first equation, I get

$$\begin{aligned}\eta \left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta} \right|_{a=1} - \left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a} \right|_{a=1} &= \frac{2(1-\eta)^2}{(1-\eta^2)^2} \left[1 + \eta + (1-\eta^2) \left(\eta u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C) - u'_{\max}(C) \frac{\partial C}{\partial a} u_{\max}(C) \right) \right] \\ &= \frac{2(1-\eta)^2}{(1-\eta^2)^2} [1 + \eta - (1-\eta^2) \lambda_{\max}(C)],\end{aligned}$$

where in the second equality I am using identity (C.31). Therefore,

$$\left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a} \right|_{a=1} = \eta \left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta} \right|_{a=1} - \frac{2(1-\eta)^2}{(1-\eta^2)^2} (1 + \eta - (1-\eta^2) \lambda_{\max}(C(1, \eta))).$$

Note that

$$\lambda_{\max}(C(1, \eta)) \leq \sum_{\tau=1}^{\infty} \eta^{\tau-1} \lambda_{\max}(C_{\tau}) < \sum_{\tau=1}^{\infty} \eta^{\tau-1} = \frac{1}{1-\eta},$$

where the inequality is by Lemma C.2. Therefore,

$$-\frac{2(1-\eta)^2}{(1-\eta^2)^2} (1 + \eta - (1-\eta^2) \lambda_{\max}(C(1, \eta))) < \frac{2(1-\eta)^2}{(1-\eta^2)^2} (1 + \eta - 1 - \eta) = 0.$$

On the other hand, by the optimality of $a^* = 1$ and $\eta^* < 1$,

$$\left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta} \right|_{a^*=1, \eta=\eta^*} \leq 0.$$

Thus,

$$\left. \frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a} \right|_{a^*=1, \eta=\eta^*} < 0,$$

a contradiction to the assumption of optimality of $a^* = 1$ and $\eta^* < 1$. This proves that $a^* < 1$ and establishes the stationarity and ergodicity of the one-state model with $a = a^*$ and $\eta = \eta^*$. \square

Proof of Theorem 5

Let λ denote the eigenvalue of C_1 largest in magnitude.⁴⁵ If $\rho(C_1) = 0$, then $\rho(C_{\tau}) = 0$ for all $\tau \geq 1$. Since C_{τ} are symmetric matrices, this implies that $C_{\tau} = 0$ for all $\tau \geq 1$. Therefore,

$$\lambda_{\max}(\Omega(a, \eta)) = -\frac{a^2(1-\eta)^2}{1-a^2\eta^2}.$$

⁴⁵The proof does not assume that λ is unique. I allow for the possibility that λ and $\lambda' = -\lambda$ are both eigenvalues of C_1 and $|\lambda| = |\lambda'| = \rho(C_1)$.

The above expression is maximized by setting $(1 - \eta)a = 0$. Therefore, by Theorem 3, for any pseudo-true one-state model, $E_t^\theta[y_{t+s}] = a^s(1 - \eta)qp' \sum_{\tau=0}^\infty a^\tau \eta^\tau y_{t-\tau} = 0$. On the other hand, if $\rho(C_1) = 0$, then $\lambda = 0$. Therefore, the theorem holds in the case $\rho(C_1) = 0$.

In the rest of the proof, I assume $\rho(C_1) > 0$. Define

$$\begin{aligned}\bar{f}(a, \eta) &\equiv -\frac{a^2(1 - \eta)^2}{1 - a^2\eta^2} + \frac{2(1 - \eta)(1 - a^2\eta)}{1 - a^2\eta^2} \sum_{\tau=1}^\infty |a|^\tau \eta^{\tau-1} \rho(C_1)^\tau \\ &= -\frac{a^2(1 - \eta)^2}{1 - a^2\eta^2} + \frac{2(1 - \eta)(1 - a^2\eta)}{1 - a^2\eta^2} \frac{|a|\rho(C_1)}{1 - \eta|a|\rho(C_1)},\end{aligned}$$

where in the second equality I am using the fact that $\rho(C_\tau) < 1$, established in Lemma C.2. Function $\bar{f}(a, \eta)$ has two maximizers given by $(\bar{a}^*, \bar{\eta}^*) = (-\rho(C_1), 0)$ and $(\bar{a}^*, \bar{\eta}^*) = (\rho(C_1), 0)$ with the maximum given by $\bar{f}^* = \rho(C_1)^2$. I establish the theorem by showing that $\lambda_{\max}(\Omega(a, \eta)) \leq \bar{f}(a, \eta)$ for all a and η , $\lambda_{\max}(\Omega(\lambda, 0)) = \bar{f}(\lambda, 0) = \bar{f}^*$, and $\lambda_{\max}(\Omega(-\lambda, 0)) \leq \bar{f}(-\lambda, 0) = \bar{f}^*$ with the inequality strict if $-\lambda$ is not an eigenvalue of C_1 . This establishes that $(a^*, \eta^*) = (\lambda, 0)$ is the unique maximizer of $\lambda_{\max}(\Omega(a, \eta))$ if $-\lambda$ is not eigenvalue of C_1 and that $(a^*, \eta^*) = (\lambda, 0)$ and $(a^*, \eta^*) = (-\lambda, 0)$ are the only maximizers of $\lambda_{\max}(\Omega(a, \eta))$ if λ and $-\lambda$ are both eigenvalues of C_1 .

As the first step in doing so, I show that for all a and τ ,

$$\lambda_{\max}(a^\tau C_\tau) \leq |a|^\tau \rho(C_1)^\tau,$$

by considering four disjoint cases: If $a \leq 0$ and $\lambda_{\min}(C_\tau) \leq 0$, then

$$\lambda_{\max}(a^\tau C_\tau) = a^\tau \lambda_{\min}(C_\tau) = |a|^\tau |\lambda_{\min}(C_\tau)| \leq |a|^\tau \rho(C_1)^\tau.$$

If $a \leq 0$ and $\lambda_{\min}(C_\tau) > 0$, then

$$\lambda_{\max}(a^\tau C_\tau) = a^\tau \lambda_{\min}(C_\tau) \leq 0 \leq |a|^\tau \rho(C_1)^\tau.$$

If $a > 0$ and $\lambda_{\max}(C_\tau) \leq 0$, then

$$\lambda_{\max}(a^\tau C_\tau) = a^\tau \lambda_{\max}(C_\tau) \leq 0 \leq |a|^\tau \rho(C_1)^\tau.$$

Finally, if $a > 0$ and $\lambda_{\max}(C_\tau) > 0$, then

$$\lambda_{\max}(a^\tau C_\tau) = a^\tau \lambda_{\max}(C_\tau) = |a|^\tau |\lambda_{\max}(C_\tau)| \leq |a|^\tau \rho(C_1)^\tau.$$

Thus, $\lambda_{\max}(a^\tau C_\tau) \leq |a|^\tau \rho(C_1)^\tau$ regardless of the value of a and the eigenvalues of C_1 . Therefore,

$$\lambda_{\max}\left(\sum_{\tau=1}^\infty a^\tau \eta^{\tau-1} C_\tau\right) \leq \sum_{\tau=1}^\infty \eta^{\tau-1} \lambda_{\max}(a^\tau C_\tau) \leq \sum_{\tau=1}^\infty \eta^{\tau-1} |a|^\tau \rho(C_1)^\tau = \frac{|a|\rho(C_1)}{1 - \eta|a|\rho(C_1)},$$

where the first inequality is using the fact that $\eta^{\tau-1} \geq 0$ for all $\tau \geq 1$ and Weyl's inequality. Consequently,

$$\lambda_{\max}(\Omega(a, \eta)) \leq \bar{f}(a, \eta) < \rho(C_1)^2$$

for any a, η such that $(|a|, \eta) \neq (\rho(C_1), 0)$.

I finish the proof by arguing that $\lambda_{\max}(\Omega(\lambda, 0)) = \rho(C_1)^2$ and $\lambda_{\max}(\Omega(-\lambda, 0)) \leq \bar{f}(-\lambda, 0) = \rho(C_1)^2$ with the inequality strict if $-\lambda$ is not an eigenvalue of C_1 . To see this, first note that

$$\lambda_{\max}(\Omega(a, 0)) = -a^2 + 2\lambda_{\max}(aC_1) = \begin{cases} -a^2 + 2a\lambda_{\min}(C_1) & \text{if } a < 0, \\ -a^2 + 2a\lambda_{\max}(C_1) & \text{if } a \geq 0. \end{cases}$$

Thus,

$$\max_{a \in [-1, 1]} \lambda_{\max}(\Omega(a, 0)) = \begin{cases} \lambda_{\min}(C_1)^2 & \text{if } |\lambda_{\min}(C_1)| > \lambda_{\max}(C_1), \\ \lambda_{\max}(C_1)^2 & \text{if } |\lambda_{\min}(C_1)| \leq \lambda_{\max}(C_1), \end{cases}$$

and

$$\arg \max_{a \in [-1, 1]} \lambda_{\max}(\Omega(a, 0)) = \begin{cases} \{\lambda_{\min}(C_1)\} & \text{if } |\lambda_{\min}(C_1)| > \lambda_{\max}(C_1), \\ \{\lambda_{\min}(C_1), \lambda_{\max}(C_1)\} & \text{if } |\lambda_{\min}(C_1)| = \lambda_{\max}(C_1), \\ \{\lambda_{\max}(C_1)\} & \text{if } |\lambda_{\min}(C_1)| < \lambda_{\max}(C_1). \end{cases}$$

Since C_1 is a symmetric matrix, the eigenvalues of C_1 are all real, and so,

$$\rho(C_1) = \begin{cases} -\lambda_{\min}(C_1) & \text{if } |\lambda_{\min}(C_1)| > \lambda_{\max}(C_1), \\ \lambda_{\max}(C_1) & \text{if } |\lambda_{\min}(C_1)| \leq \lambda_{\max}(C_1). \end{cases}$$

This establishes that, in any pseudo-true one-state model, $\eta = 0$, $a = \lambda$, and

$$\Omega(a, \eta) = -\lambda^2 I + 2\lambda C_1.$$

By Theorem 3, u is an eigenvector of $\Omega(a, \eta)$ with eigenvalue $\lambda_{\max}(\Omega(a, \eta)) = \lambda^2$ and $u'u = 1$. Therefore, u is also an eigenvector of C_1 , but with eigenvalue λ . This completes the proof of the theorem. \square

Proof of Theorem 6

I first prove a useful lemma, which offers a canonical representation of matrices C_l in the case where the true process can be represented as in (6):⁴⁶

Lemma C.3. *Suppose $\{C_l\}_l$ are the autocorrelation matrices of an n -dimensional stationary ergodic process that can be represented as in (6) with $f_t \in \mathbb{R}^m$. There exists a convergent $m \times m$ matrix \mathbb{F} with $\|\mathbb{F}\|_2 \leq 1$, and a semi-orthogonal $m \times n$ matrix \mathbb{H} such that*

$$C_l = \mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H}. \quad (\text{C.32})$$

Conversely, for any positive integers $m \geq n$, $m \times m$ convergent matrix \mathbb{F} with $\|\mathbb{F}\|_2 \leq 1$, and semi-orthogonal $m \times n$ matrix \mathbb{H} , there exists an n -dimensional stationary ergodic process with autocorrelation matrices $\{C_l\}_l$ of the form (C.32), which can be represented as in (6).⁴⁷

⁴⁶Versions of this result have previously appeared in the control and time-series literatures. For early examples, see Ho and Kálmán (1966) and Akaike (1975).

⁴⁷Matrix $\mathbb{H} \in \mathbb{R}^{m \times n}$ is semi-orthogonal if $\mathbb{H}'\mathbb{H} = I$, where I denotes the $n \times n$ identity matrix.

Proof. Recall that I have assumed (without loss of generality) that the true process is non-degenerate, i.e., $\mathbb{E}[y_t y_t']$ is invertible. Invertibility of $\mathbb{E}[y_t y_t']$ requires $m \geq n$, an assumption I maintain throughout the first part of the proof. Given representation (6), the autocovariance matrices are given by

$$\Gamma_l = \mathbb{E}[y_t y_{t-l}'] = H' F^l \mathbb{E}[f_t f_t'] H = H' F^l V H,$$

where $V \equiv \mathbb{E}[f_t f_t']$ is the unique solution to the following discrete-time Lyapunov equation:

$$V = F V F' + \Sigma, \quad (\text{C.33})$$

and Σ is the variance-covariance matrix of ϵ_t . Therefore,

$$C_l = (H' V H)^{\frac{-1}{2}} \left(\frac{H' F^l V H + H' V F'^l H}{2} \right) (H' V H)^{\frac{-1}{2}}.$$

Matrix V is positive semidefinite; it is positive definite if the representation in (6) is minimal.⁴⁸ Without loss of generality, I assume that that is the case. Define

$$\begin{aligned} \mathbb{H}' &\equiv (H' V H)^{\frac{-1}{2}} H' V^{\frac{1}{2}}, \\ \mathbb{F} &\equiv V^{\frac{-1}{2}} F V^{\frac{1}{2}}. \end{aligned}$$

Then

$$C_l = \mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H}. \quad (\text{C.34})$$

Note that since F is a convergent matrix, so is \mathbb{F} . Substituting $\mathbb{F} = V^{\frac{-1}{2}} F V^{\frac{1}{2}}$ in equation (C.33), I get

$$I - \mathbb{F} \mathbb{F}' = V^{\frac{-1}{2}} \Sigma V^{\frac{-1}{2}}.$$

Therefore, since Σ is positive semidefinite, the spectral radius of $\mathbb{F} \mathbb{F}'$ is weakly smaller than one. This implies that $\|\mathbb{F}\|_2 \leq 1$. On the other hand,

$$\mathbb{H}' \mathbb{H} = (H' V H)^{\frac{-1}{2}} H' V H (H' V H)^{\frac{-1}{2}} = I.$$

That is, \mathbb{H} is a (full-rank) semi-orthogonal matrix. This proves the first part of the theorem.

I next argue that given a convergent matrix $\hat{\mathbb{F}} \in \mathbb{R}^{m \times m}$ with $\|\hat{\mathbb{F}}\|_2 \leq 1$ and a semi-orthogonal matrix $\hat{\mathbb{H}} \in \mathbb{R}^{m \times n}$ with $m \geq n$, there exists a stationary ergodic process such that the corresponding autocorrelation matrices are given by (C.34) with $\mathbb{F} = \hat{\mathbb{F}}$ and $\mathbb{H} = \hat{\mathbb{H}}$. Given any such $\hat{\mathbb{F}}$ and $\hat{\mathbb{H}}$, let $F = \hat{\mathbb{F}}$, $H = \hat{\mathbb{H}}$, and $\Sigma = I - \hat{\mathbb{F}} \hat{\mathbb{F}}'$. The solution to the Lyapunov equation (C.33) is then given by $V = I$. Therefore, $\mathbb{F} = F = \hat{\mathbb{F}}$ and $\mathbb{H} = \hat{\mathbb{H}}(\hat{\mathbb{H}}' \hat{\mathbb{H}})^{\frac{-1}{2}} = \hat{\mathbb{H}}$, where in the last equality I am using the assumption of semi-orthogonality of $\hat{\mathbb{H}}$. By construction, then the autocorrelation matrices of a process of the form (6) with matrices F , H , and Σ as above are given by (C.34) with $\mathbb{F} = \hat{\mathbb{F}}$ and $\mathbb{H} = \hat{\mathbb{H}}$. \square

⁴⁸See, for instance, Akaike (1975).

Proof of Theorem 6. I assume without loss of generality that the representation in (6) is minimal. By Lemma C.3 then,

$$C_l = \mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H},$$

where $\mathbb{H}' \equiv (H'VH)^{-\frac{1}{2}} H'V^{\frac{1}{2}}$, $\mathbb{F} \equiv V^{-\frac{1}{2}} FV^{\frac{1}{2}}$, and $V \equiv \mathbb{E}[f_t f_t']$ is the variance-covariance of f_t . Note that since the variance-covariance of f_t is normalized to be the identity matrix, $V = I$, $\mathbb{F} = F$, and $\mathbb{H} = H$. Recall that vector y_t does not contain any redundant observables (which are linear combinations of other observables). This assumption, together with the assumption that H is a rank- m matrix, ensures that H is an invertible $m \times m$ matrix. Therefore, by Lemma C.3, $\mathbb{H} = H$ is an orthogonal matrix. Thus,

$$\rho(C_l) = \rho \left(\mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H} \right) = \rho \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) = \rho \left(\frac{F^l + F'^l}{2} \right) \quad (\text{C.35})$$

for all l . But since the spectral radius of a symmetric matrix equals its spectral norm,

$$\rho \left(\frac{F^l + F'^l}{2} \right) = \left\| \frac{F^l + F'^l}{2} \right\|_2 \leq \frac{1}{2} \|F^l\|_2 + \frac{1}{2} \|F'^l\|_2 = \|F^l\|_2 \leq \|F\|_2^l. \quad (\text{C.36})$$

Therefore,

$$\rho(C_l) \leq \|F\|_2^l.$$

On the other hand, by equations (C.35) and (C.36),

$$\rho(C_1) = \left\| \frac{F + F'}{2} \right\|_2 = \|F\|_2,$$

where the second equality is by assumption. Thus,

$$\rho(C_l) \leq \|F\|_2^l = \rho(C_1)^l,$$

and the process is exponentially ergodic. \square

Proof of Proposition 1

I first state and prove a useful lemma:

Lemma C.4. Suppose C_1 has a unique and simple eigenvalue λ with $|\lambda| = \rho(C_1) > 0$, and let u denote the corresponding eigenvector normalized to have $u'u = 1$.⁴⁹ If $u'C_2u > \rho(C_1)^2$, then the agent's forecasts in any pseudo-true one-state model are given by (4) with a tuple (a, η, p, q) such that $\eta > 0$.

Proof. Define $C(a, \eta)$ as in the proof of Theorem 4. As in the proof of Theorem 4, I present the argument under the assumption that the largest eigenvalue of $C(a, \eta)$ is simple at the point (a^*, η^*) that maximizes $\lambda_{\max}(C(a, \eta))$.⁵⁰ I start by proposing a candidate solution to the problem

⁴⁹The assumption that λ is unique and simple is not necessary for the result. The result generalizes to arbitrary matrices C_1 with $\rho(C_1) \neq 0$ by replacing $u'C_2u$ with the maximum of $u'C_2u$ over all unit-norm eigenvectors u of C_1 with eigenvalues λ such that $|\lambda| = \rho(C_1)$.

⁵⁰See footnote 43 for how the argument can be generalized.

of maximizing $\lambda_{\max}(\Omega(a, \eta))$ at which $\eta = 0$ and argue that the candidate does not satisfy the necessary first-order optimality conditions. Setting $\eta = 0$ in equations (C.29) and (C.30), I get

$$\begin{aligned}\left.\frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a}\right|_{\eta=0} &= -2a + 2u'_{\max}(aC_1)C_1u_{\max}(aC_1), \\ \left.\frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta}\right|_{\eta=0} &= 2a^2 - 2(1 + a^2)\lambda_{\max}(aC_1) + 2a^2u'_{\max}(aC_1)C_2u_{\max}(aC_1),\end{aligned}$$

where I am using the fact that $C = aC_1$ when $\eta = 0$. Any solution to $\partial \lambda_{\max}(\Omega(a, \eta))/\partial a|_{\eta=0} = 0$ satisfies $a = \lambda$, where $\lambda = \lambda_{\min}(C_1)$ if $\lambda_{\max}(C_1) \leq 0$, $\lambda = \lambda_{\max}(C_1)$ if $\lambda_{\min}(C_1) \geq 0$, and $\lambda \in \{\lambda_{\max}(C_1), \lambda_{\min}(C_1)\}$ otherwise. Evaluating $\lambda_{\max}(\Omega(a, \eta))$ at $a = \lambda$ and $\eta = 0$, I get $\lambda_{\max}(\Omega(\lambda, 0)) = \lambda^2$. Therefore, for the solution $(a, \eta) = (\lambda, 0)$ to the first-order condition $\partial \lambda_{\max}(\Omega(a, \eta))/\partial a = 0$ to be a maximizer of $\lambda_{\max}(\Omega(a, \eta))$, it must be the case that λ is the eigenvalue of C_1 largest in magnitude and $u = u_{\max}(aC_1)$ is a corresponding eigenvector normalized such that $u'u = 1$. Substituting in the expression for $\partial \lambda_{\max}(\Omega(a, \eta))/\partial \eta|_{\eta=0}$, I get

$$\left.\frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial \eta}\right|_{a=\lambda, \eta=0} = 2\rho(C_1)^2 (u'C_2u - \rho(C_1)^2) > 0,$$

where the inequality follows the assumption that $u'C_2u > \rho(C_1)^2$. This implies that the pair $\eta = 0$ and $a = \lambda$ does not constitute a local maximizer of $\lambda_{\max}(\Omega(a, \eta))$. Since this pair is the only candidate with $\eta = 0$ that may satisfy the first-order conditions, in any pseudo-true one-state model, $\eta > 0$. This establishes the lemma. \square

Proof of Proposition 1. Let σ^2 denote the variance of y_t . By the argument in the proof of Lemma C.3, the lag- l autocorrelation of y_t is given by

$$C_l = H' \left(\frac{F^l + F'^l}{2} \right) H,$$

where $F = V^{\frac{1}{2}} \mathbb{F} V^{\frac{1}{2}}$, $H' = (\mathbb{H}' V \mathbb{H})^{\frac{1}{2}} \mathbb{H}' V^{\frac{1}{2}}$, and V is the solution to the discrete-time Lyapunov equation (C.33). Since \mathbb{F} and Σ are diagonal matrices, so is V . Therefore, $F = \mathbb{F}$. On the other hand, by Lemma C.3, H is a semi-orthogonal matrix. Therefore, $H'H = 1$, and so,

$$C_l = \sum_{i=1}^m w_i \alpha_i^l,$$

where $w_i = H_i^2 \geq 0$, $\sum_{i=1}^m w_i = 1$, and α_i is the i th diagonal element of \mathbb{F} . That is, $C_l^{\frac{1}{l}}$ is equal to the weighted l -norm of vector $(\alpha_1, \dots, \alpha_m)$ with weights $w = (w_1, \dots, w_m)$.

Since the representation in (6) is minimal, $w_i > 0$ for all i , and all α_i are distinct. If that were not the case, there would exist some $\tilde{m} < m$ such that $C_l = \sum_{i=1}^{\tilde{m}} \tilde{w}_i \tilde{\alpha}_i^l$ for some non-negative weights \tilde{w}_i that sum up to one and some $\tilde{\alpha}_i \in (-1, 1)$. Consider the process $\tilde{\mathbb{P}}$ represented as in (6) with $\mathbb{F} = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_{\tilde{m}})$, $\epsilon_t \sim \mathcal{N}(0, \Sigma)$, $\Sigma = I - \mathbb{F}\mathbb{F}'$, and $\mathbb{H} = \sigma \text{diag}(\sqrt{\tilde{w}_1}, \dots, \sqrt{\tilde{w}_{\tilde{m}}})$. By the argument in the proof of Lemma C.3, $\tilde{\mathbb{P}}$ has the same autocorrelation matrices as \mathbb{P} . Moreover,

both \mathbb{P} and $\tilde{\mathbb{P}}$ are mean-zero and normal and both have variance σ^2 . Therefore, \mathbb{P} and $\tilde{\mathbb{P}}$ are observationally equivalent, a contradiction to the assumption that the representation I started with was minimal.

Next, note that, by the generalized mean inequality, $C_l^{\frac{1}{l}} > C_1$ for all $l \geq 2$, where the strictness of the inequality follows the facts that $w_i > 0$ for all i and all α_i are distinct. In particular, $u' C_2 u = C_2 > C_1^2 = \rho(C_1)^2$, where I am using the fact that y_t is a scalar. Thus, by Lemma C.4, $\eta > 0$. To see why $\eta < 1$, recall that by Theorem 3, the (a, η) pair maximizes

$$\Omega(\tilde{a}, \tilde{\eta}) = -\frac{\tilde{a}^2(1 - \tilde{\eta})^2}{1 - \tilde{a}^2\tilde{\eta}^2} + \frac{2(1 - \tilde{\eta})(1 - \tilde{a}^2\tilde{\eta})}{1 - \tilde{a}^2\tilde{\eta}^2} \sum_{\tau=1}^{\infty} \tilde{a}^{\tau} \tilde{\eta}^{\tau-1} C_{\tau}.$$

But $\Omega(\tilde{a}, 1) = 0 < C_1^2 = \Omega(C_1, 0)$ for any \tilde{a} . Therefore, $\eta = 1$ cannot be part of the description of a pseudo-true one-state model. Finally, $a \in (1, 1)$ by Theorem 4. The proposition then follows Theorem 3 by noting that $qp' = 1$ whenever y_t is a scalar. \square

Proof of Theorem 7

I first find a transformation $\tilde{y}_t = T y_t$ of the vector of observables such that T is invertible and $\tilde{\Gamma}_0^{-\frac{1}{2}} \tilde{\Gamma}_1 \tilde{\Gamma}_0^{-\frac{1}{2}}$ is diagonal. Since matrices Γ_0 and Γ_1 are both symmetric and Γ_0 is non-singular, Γ_0 and Γ_1 can be diagonalized simultaneously by a real congruence transformation. Note that since Γ_1 is symmetric,

$$\Gamma_0^{-\frac{1}{2}} \Gamma_1 \Gamma_0^{-\frac{1}{2}} = \frac{1}{2} \Gamma_0^{-\frac{1}{2}} (\Gamma_1 + \Gamma_1') \Gamma_0^{-\frac{1}{2}} = C_1$$

is symmetric. Therefore, there exists a diagonal matrix Λ , with the eigenvalues of C_1 as its diagonal elements, and an orthogonal matrix U such that $C_1 = U \Lambda U'$. Define

$$T \equiv U' \Gamma_0^{-\frac{1}{2}}.$$

It is easy to verify that $T \Gamma_0 T' = I$ and $T \Gamma_1 T' = \Lambda$. The autocovariance matrices of $\tilde{y}_t \equiv T y_t$ are given by $\tilde{\Gamma}_l \equiv \mathbb{E}[\tilde{y}_t \tilde{y}_{t-l}'] = T \Gamma_l T'$. In particular, $\tilde{\Gamma}_0 = I$, $\tilde{\Gamma}_1 = \Lambda$, and so $\tilde{\Gamma}_0^{-\frac{1}{2}} \tilde{\Gamma}_1 \tilde{\Gamma}_0^{-\frac{1}{2}} = \Lambda$.

I find the pseudo-true m.i.o. d -state models when the vector of observables is \tilde{y}_t and then transform the models back using the linear-invariance result to find the pseudo-true m.i.o. d -state models when the observable is y_t . By Lemma C.1, the KLDR and the agents' forecasts can be represented in terms of matrices \tilde{M} , \tilde{N} , and \tilde{D} of the transformed model as in (C.9) and (C.16). Let $S \equiv \tilde{\Gamma}_0^{\frac{1}{2}} \tilde{N} = \tilde{N}$ and use the restriction to the set of m.i.o. models to set $\tilde{D} = D = (I \ \mathbf{0})'$. The expression for the KLDR in (C.9) then simplifies to

$$\text{KLDR}(\theta) = -\frac{1}{2} \log \det(SS') + \frac{1}{2} \text{tr}(S'S) - \text{tr}(\tilde{M} D' S' \Lambda S D) + \frac{1}{2} \text{tr}(\tilde{M} D' S' S D \tilde{M}') + \text{constant}.$$

I begin by ignoring the constraints that \tilde{M} is a convergent matrix and $\|\tilde{M}(I - D'D)\tilde{M}'\|_2 < 1$ and showing that the solution to the relaxed problem is also a solution to the original problem. For the relaxed problem, the necessary first-order optimality conditions with respect to S and \tilde{M} are

given by

$$I = S'S - S'\Lambda'S\tilde{D}\tilde{M}'D' - S'\Lambda S\tilde{D}\tilde{M}'D' + S'S\tilde{D}\tilde{M}'\tilde{M}D', \quad (\text{C.37})$$

$$D'S'\Lambda S\tilde{D} = \tilde{M}'D'S'S\tilde{D}. \quad (\text{C.38})$$

I proceed by first characterizing the set of all solutions for S and \tilde{M} to these optimality conditions. Given any such solution, the KLDR is given by $-\log \det (SS')/2 + \text{constant}$. Therefore, if there are multiple solutions to equations (C.37) and (C.38), the optimal solution is the one with the largest value of $\log \det (SS')$.

In the next step, I write $S = (S_1 \ S_2)$ where $S_1 \in \mathbb{R}^{n \times d}$ and $S_2 \in \mathbb{R}^{n \times (n-d)}$. Equation (C.37) then can be written as

$$\begin{aligned} \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} &= \begin{pmatrix} S_1'S_1 & S_1'S_2 \\ S_2'S_1 & S_2'S_2 \end{pmatrix} - \begin{pmatrix} S_1'\Lambda'S_1 & S_1'\Lambda'S_2 \\ S_2'\Lambda'S_1 & S_2'\Lambda'S_2 \end{pmatrix} \begin{pmatrix} \tilde{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} - \begin{pmatrix} S_1'\Lambda S_1 & S_1'\Lambda S_2 \\ S_2'\Lambda S_1 & S_2'\Lambda S_2 \end{pmatrix} \begin{pmatrix} \tilde{M}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \\ &\quad + \begin{pmatrix} S_1'S_1 & S_1'S_2 \\ S_2'S_1 & S_2'S_2 \end{pmatrix} \begin{pmatrix} \tilde{M}'\tilde{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} S_1'S_1 & S_1'S_2 \\ S_2'S_1 & S_2'S_2 \end{pmatrix} - \begin{pmatrix} S_1'\Lambda'S_1\tilde{M} & \mathbf{0} \\ S_2'\Lambda'S_1\tilde{M} & \mathbf{0} \end{pmatrix} - \begin{pmatrix} S_1'\Lambda S_1\tilde{M}' & \mathbf{0} \\ S_2'\Lambda S_1\tilde{M}' & \mathbf{0} \end{pmatrix} + \begin{pmatrix} S_1'S_1\tilde{M}'\tilde{M} & \mathbf{0} \\ S_2'S_1\tilde{M}'\tilde{M} & \mathbf{0} \end{pmatrix}. \end{aligned}$$

Equivalently,

$$S_1'S_1 - S_1'\Lambda'S_1\tilde{M} - S_1'\Lambda S_1\tilde{M}' + S_1'S_1\tilde{M}'\tilde{M} = I, \quad (\text{C.39})$$

$$S_2'S_1 = \mathbf{0}, \quad (\text{C.40})$$

$$S_2'\Lambda'S_1\tilde{M} + S_2'\Lambda S_1\tilde{M}' = \mathbf{0}, \quad (\text{C.41})$$

$$S_2'S_2 = I. \quad (\text{C.42})$$

Likewise, equation (C.38) can be written as

$$S_1'\Lambda S_1 = \tilde{M}'S_1'S_1. \quad (\text{C.43})$$

Equation (C.42) implies that S_2 is a full-rank matrix. On the other hand, since $S'S = \begin{pmatrix} S_1'S_1 & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix}$ is invertible, S_1 is also a full-rank matrix. Equation (C.40) then implies that the column space (or range) of S_1 is the same as the null space (or kernel) of S_2' , and equation (C.41) implies that the column space of $\Lambda'S_1\tilde{M} + \Lambda S_1\tilde{M}'$ is a subspace of the null space of S_2' . Therefore,

$$\Lambda'S_1\tilde{M} + \Lambda S_1\tilde{M}' = S_1Y \quad (\text{C.44})$$

for some matrix Y .

Define $X \equiv S_1'S_1$ and $Z \equiv S_1'\Lambda S_1$. Since S_1 is full rank, X is invertible. Left-multiplying equation (C.44) by S_1' , I get

$$Z'\tilde{M} + Z\tilde{M}' = XY. \quad (\text{C.45})$$

Meanwhile, equations (C.39) and (C.43) can be written as

$$X - Z'\tilde{M} - Z\tilde{M}' + X\tilde{M}'\tilde{M} = I, \quad (\text{C.46})$$

$$Z = \tilde{M}X. \quad (\text{C.47})$$

Solving for \tilde{M} from (C.47) and substituting in (C.46), I get

$$X = I + ZX^{-1}Z'. \quad (\text{C.48})$$

Therefore, X is a symmetric positive-definite matrix with eigenvalues weakly larger than one. Combining (C.45) and (C.46), I get

$$X - XY + X\tilde{M}'\tilde{M} = I,$$

and so

$$Y = I + X^{-1}Z'ZX^{-1} - X^{-1}.$$

Because the eigenvalues of X are weakly larger than one, Y is a symmetric positive semi-definite matrix.

Let $\mathcal{K} \subseteq \mathbb{R}^d$ denote the (possibly empty) null space of $Z' = Z$. If $p \in \mathcal{K}$, then

$$Xp = p + ZX^{-1}Z'p = p,$$

$$X^{-1}p = p,$$

and

$$Yp = p + X^{-1}Z'ZX^{-1}p - X^{-1}p = X^{-1}Z'Zp = 0,$$

where in the last equality I am using the symmetry of Z . Therefore, \mathcal{K} is an eigenspace of both X and X^{-1} with eigenvalue one and a subspace of the null space of Y . It is easy to show that, if p is an eigenvector of X with eigenvalue one, then it is also in the null spaces of $Z = Z'$ and Y , and that if p is in the null space of Y , it is also in the null space of $Z = Z'$ and an eigenvector of X with eigenvalue one. Therefore, \mathcal{K} is identical to the null space of Y and the eigenspace of X with eigenvalue one. By a similar argument, equations (C.45) and (C.47) imply that \mathcal{K} is also the null space of $\tilde{M} + \tilde{M}'$.

I next show that the column space of S_1 is spanned by d eigenvectors of Λ . Decompose \mathbb{R}^d as a direct sum: $\mathbb{R}^d = \mathcal{K} \oplus \mathcal{R}$, where \mathcal{K} is the kernel of $Z' = Z$ and \mathcal{R} is the range of Z . Let d_K denote the dimension of \mathcal{K} and d_R denote the dimension of \mathcal{R} . By the rank-nullity theorem, $d_K + d_R = d$. Any vector $p \in \mathbb{R}^d$ can be written as $p = p_K + p_R$, where $p_K \in \mathcal{K}$ and $p_R \in \mathcal{R}$. First, consider the set of $S_1 p_K$ for $p_K \in \mathcal{K}$. By construction, $p_K S_1' \Lambda S_1 p_K = 0$ for any $p_K \in \mathcal{K}$. Therefore, since S_1 is a full-rank matrix, $S_1 p_K$ is in the null space of Λ . But Λ is a diagonal matrix with its null space the same as its (possibly empty) eigenspace with eigenvalue zero. Therefore, the set $\{S_1 p_K : p_K \in \mathcal{K}\}$ is a d_K -dimensional invariant subspace of Λ spanned by eigenvectors of Λ with eigenvalue zero. Next consider the set of $S_1 p_R$ for $p_R \in \mathcal{R}$. By the argument in the previous paragraph, matrices Y and $\tilde{M} + \tilde{M}'$ are invertible when restricted to subspace \mathcal{R} with the inverse over the subspace given by the Moore–Penrose pseudo-inverse. Therefore, equation (C.44) can be written as

$$\Lambda S_1 = S_1 Y (\tilde{M} + \tilde{M}')^\dagger$$

over this subspace. But since Y and $(\tilde{M} + \tilde{M}')^\dagger$ are invertible when restricted to \mathcal{R} , the above equation implies that

$$\{S_1 p_R : p_R \in \mathcal{R}\} = \{S_1 Y (\tilde{M} + \tilde{M}')^\dagger p_R : p_R \in \mathcal{R}\} = \{\Lambda S_1 p_R : p_R \in \mathcal{R}\}.$$

That is, set $\{S_1 p_R : p_R \in \mathcal{R}\}$ is a d_R -dimensional invariant subspace of Λ . But the invariant subspaces of Λ are spanned by its eigenvectors.

The eigenvectors of Λ are the n coordinate vectors. Without loss of generality, assume that the elements of \tilde{y}_t are ordered such that the column space of S_1 is spanned by the first d coordinate vectors e_1, \dots, e_d . In such coordinates, S_1 can be written as

$$S_1 = \begin{pmatrix} S_{11} \\ \mathbf{0} \end{pmatrix},$$

where $S_{11} \in \mathbb{R}^{d \times d}$ is an invertible matrix. Therefore,

$$X = S'_1 S_1 = S'_{11} S_{11},$$

$$Z = S'_{11} \Lambda_1 S_{11},$$

where $\Lambda = \begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \Lambda_2 \end{pmatrix}$, and Λ_1 is a $d \times d$ diagonal matrix. Substituting in (C.48), I get

$$S'_{11} S_{11} = I + S'_{11} \Lambda_1 S_{11} (S'_{11} S_{11})^{-1} S'_{11} \Lambda_1 S_{11} = I + S'_{11} \Lambda_1^2 S_{11},$$

and so,

$$S'_{11} (I - \Lambda_1^2) S_{11} = I.$$

Multiplying the above equation from left by S_{11} and from right by $S_{11}^{-1} (I - \Lambda_1^2)^{-1}$, I get

$$S_{11} S'_{11} = (I - \Lambda_1^2)^{-1}. \quad (\text{C.49})$$

Therefore,

$$\log \det (SS') = \log \det (S'S) = \log \det (S'_1 S_1) = \log \det (S'_{11} S_{11}) = \log \det (S_{11} S'_{11}) = -\log \det (I - \Lambda_1^2).$$

Thus, minimizing the KLDR requires the column space of S_1 to be spanned by eigenvectors p_1, \dots, p_d of Λ corresponding to the d eigenvalues of Λ largest in magnitude.

It only remains to compute matrices $\tilde{A}, \tilde{B}, \tilde{Q}, \tilde{R}$. Note that matrix S can be written as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ \mathbf{0} & S_{22} \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} X & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} = S'S = \begin{pmatrix} S'_{11} S_{11} & S'_{11} S_{12} \\ S'_{12} S_{11} & S'_{12} S_{12} + S'_{22} S_{22} \end{pmatrix}.$$

Since S_{11} is an invertible matrix, the above equation implies that $S_{12} = \mathbf{0}$ and $S'_{22} S_{22} = I$. Therefore, since S_{22} is a symmetric invertible matrix, $S'_{22} = S_{22}^{-1}$. I can now find $\tilde{A}, \tilde{B}, \tilde{Q}, \tilde{R}$:

$$\tilde{A} = \tilde{M} = S'_{11} \Lambda_1 S_{11}^{-1},$$

$$\begin{aligned}\tilde{B} &= D' S^{-1} \tilde{\Gamma}_0^{\frac{1}{2}} = (S_{11}^{-1} \quad \mathbf{0}), \\ \tilde{Q} &= I - \tilde{M} (I - D' D) \tilde{M}' = I, \\ \tilde{R} &= S'^{-1} (I - D D') S^{-1} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix}.\end{aligned}$$

Note that the constraints that \tilde{M} is convergent $\|\tilde{M} (I - D' D) \tilde{M}'\|_2 < 1$ are automatically satisfied. I can use the above expressions to compute the forecasts of \tilde{y}_t . Equation (C.16) implies

$$E_t^\theta [\tilde{y}_{t+s}] = \begin{pmatrix} S_{11}'^{-1} \\ \mathbf{0} \end{pmatrix} S_{11}' \Lambda_1^s S_{11}'^{-1} (S_{11} \quad \mathbf{0}) \tilde{y}_t = \begin{pmatrix} \Lambda_1^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tilde{y}_t.$$

Using $\tilde{y}_t = T y_t$ for all t in the above equation, I get

$$E_t^\theta [y_{t+s}] = T^{-1} \begin{pmatrix} \Lambda_1^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} T y_t = \Gamma_0^{\frac{1}{2}} U \begin{pmatrix} \Lambda_1^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} U' \Gamma_0^{-\frac{1}{2}} y_t = \Gamma_0^{\frac{1}{2}} U D' U' C_1^s U D U' \Gamma_0^{-\frac{1}{2}} y_t.$$

Using the definition of U , I can simplify the above expression to

$$E_t^\theta [y_{t+s}] = \sum_{i=1}^d \Gamma_0^{\frac{1}{2}} u_i \lambda_i^s u_i' \Gamma_0^{-\frac{1}{2}} y_t,$$

where λ_i is the i th largest eigenvalue of C_1 , and u_i is the corresponding eigenvector normalized such that $u_i' u_i = 1$. The theorem then follows the definitions of a_i , p_i , and q_i . \square

Proof of Theorem 8

Setting $M = a$, $D = \sqrt{1 - \eta} e_1$, and $N = \Gamma_0^{-\frac{1}{2}} S$ in equation (C.17), I get

$$\text{Var}^\theta(y) = \Gamma_0^{\frac{1}{2}} \left[I + \frac{1}{1 - a^2} \left[a^2(1 - \eta)^2 - (1 - 2a^2\eta + a^2\eta^2) \lambda \right] u u' \right] \Gamma_0^{\frac{1}{2}},$$

where a , η , $\lambda = \lambda_{\max}(\Omega(a, \eta))$, and u are as in Theorem 3. Substituting for $\lambda_{\max}(\Omega(a, \eta))$ from equation (C.28) in the above equation, I get

$$\text{Var}^{1*}(y_t) = \Gamma_0^{\frac{1}{2}} \left[I + \frac{2(1 - \eta)(1 - a^2\eta)}{(1 - a^2)(1 - a^2\eta^2)} \left(a^2(1 - \eta) - (1 - 2a^2\eta + a^2\eta^2) \lambda_{\max}(C) \right) u u' \right] \Gamma_0^{\frac{1}{2}}. \quad (\text{C.50})$$

Let a^* and η^* be scalars in the $[-1, 1]$ and $[0, 1]$ intervals, respectively, that maximize $\lambda_{\max}(\Omega(a, \eta))$. I separately consider the cases $\eta^* = 1$ and $\eta^* < 1$. If $\eta^* = 1$, then the right-hand side of equation (C.50) is equal to Γ_0 .

Next suppose $\eta^* < 1$. By the argument in the proof of Theorem 4, the first-order optimality condition with respect to a must hold with equality at $a = a^*$ and $\eta = \eta^* < 1$. Setting $\partial \lambda_{\max}(\Omega(a, \eta)) / \partial a = 0$ in (C.29) and multiplying both sides of the equation by a^* , I get, using (C.31),

$$\begin{aligned}& \frac{2a^{*2}(1 - \eta^*)^2}{(1 - a^{*2}\eta^{*2})^2} + \frac{4a^{*2}\eta^*(1 - \eta^*)^2}{(1 - a^{*2}\eta^{*2})^2} \lambda_{\max}(C) \\ &= \frac{2(1 - \eta^*)(1 - a^{*2}\eta^*)}{1 - a^{*2}\eta^{*2}} \lambda_{\max}(C) + \frac{2(1 - \eta^*)(1 - a^{*2}\eta^*)}{1 - a^{*2}\eta^{*2}} \eta^* u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C).\end{aligned} \quad (\text{C.51})$$

Setting $\eta^* = 0$ in the above equation, I get $a^{*2} = \lambda_{\max}(C)$. Setting $a^{*2} = \lambda_{\max}(C)$ in equation (C.50) then establishes that $\text{Var}^{1*}(y_t) = \Gamma_0$ in the case where $\eta^* = 0$.

Finally, I consider the case where $\eta^* \in (0, 1)$. Then additionally the first-order optimality condition with respect to η must hold with equality. Setting $\partial \lambda_{\max}(\Omega(a, \eta)) / \partial \eta = 0$ in equation (C.30), multiplying it by η^* , solving for $\eta^* u'_{\max}(C) \frac{\partial C}{\partial \eta} u_{\max}(C)$, and substituting in equation (C.51), I get

$$\begin{aligned} & \frac{2a^{*2}(1-\eta^*)^2}{(1-a^{*2}\eta^{*2})^2} + \frac{4a^{*2}\eta^*(1-\eta^*)^2}{(1-a^{*2}\eta^{*2})^2} \lambda_{\max}(C) \\ &= \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}} \lambda_{\max}(C) - \frac{2a^{*2}\eta^*(1-\eta^*)(1-a^{*2}\eta^*)}{(1-a^{*2}\eta^{*2})^2} \\ & \quad + \frac{2\eta^* \left(1 + a^{*4}\eta^{*2} + a^{*2}(1-4\eta^* + \eta^{*2})\right)}{(1-a^{*2}\eta^{*2})^2} \lambda_{\max}(C). \end{aligned}$$

Simplifying the above expression leads to

$$a^{*2}(1-\eta^*) = \left(1 - 2a^{*2}\eta^* + a^{*2}\eta^{*2}\right) \lambda_{\max}(C).$$

Combining the above identity with equation (C.50) implies that $\text{Var}^{1*}(y_t) = \Gamma_0$ and finishes the proof of the theorem. \square

Proof of Theorem 9

Define T as in the proof of Theorem 7 and let $\tilde{y}_t = Ty_t$. Then by the argument in the proof of Theorem 7, $\tilde{\Gamma}_0 = I$ and $\tilde{\Gamma}_1 = \Lambda$, where Λ is a diagonal matrix with the eigenvalues of C_1 as its diagonals. Furthermore, any pseudo-true m.i.o. d -state model $\tilde{\theta} = (\tilde{A}, \tilde{B}, \tilde{Q}, \tilde{R})$ given observable \tilde{y}_t satisfies

$$\begin{aligned} \tilde{A} &= S'_{11} \Lambda_1 S'_{11}{}^{-1}, \\ \tilde{B} &= (S_{11}^{-1} \quad \mathbf{0}), \\ \tilde{Q} &= I, \\ \tilde{R} &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix}, \end{aligned}$$

where S_{11} is $d \times d$ matrix that satisfies $S_{11} S'_{11} = (I - \Lambda_1^2)^{-1}$, and Λ_1 is a $d \times d$ diagonal matrix containing the d largest eigenvalues of C_1 . On the other hand, equation (C.17) implies

$$\begin{aligned} \text{Var}^{d*}(\tilde{y}_t) &= \tilde{B}' \sum_{\tau=0}^{\infty} \tilde{A}^{\tau} \tilde{Q} (\tilde{A}')^{\tau} \tilde{B} + \tilde{R} = \left(S'_{11} \Lambda_1^{\tau} (S_{11} S'_{11})^{-1} \Lambda_1^{\tau} S_{11} (S_{11}^{-1} \quad \mathbf{0}) + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} \right) \\ &= \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} = I. \end{aligned}$$

Therefore, $\text{Var}^{d*}(y) = T^{-1} \text{Var}^{d*}(\tilde{y}_t) T'^{-1} = (T'T)^{-1} = \Gamma_0$. \square

Proof of Proposition 2

Recall that I assumed (without loss of generality) that Γ_0 is non-singular. Since $\{u_i\}_{i=1}^d$ constitutes an orthonormal basis for \mathbb{R}^n , $\Gamma_0^{-\frac{1}{2}} y_t$ can be expressed as

$$\Gamma_0^{-\frac{1}{2}} y_t = \sum_{i=1}^n \omega_{it} u_i,$$

where $\omega_{it} \equiv u_i' \Gamma_0^{-\frac{1}{2}} y_t$. Therefore,

$$y_t = \Gamma_0^{\frac{1}{2}} \sum_{i=1}^n \omega_{it} u_i = \sum_{i=1}^n \Gamma_0^{\frac{1}{2}} u_i u_i' \Gamma_0^{-\frac{1}{2}} y_t = \sum_{i=1}^n y_t^{(i)} q_i,$$

where the last equality uses the definitions of $y_t^{(i)}$ and q_i .

The lag-one autocovariance of $y_t^{(i)}$ is given by

$$\mathbb{E} \left[y_t^{(i)} y_{t-1}^{(i)} \right] = p_i' \mathbb{E} [y_t y_{t-1}] p_i = p_i' \Gamma_1 p_i = p_i' \left(\frac{\Gamma_1 + \Gamma_1'}{2} \right) p_i = p_i' \Gamma_0^{\frac{1}{2}} C_1 \Gamma_0^{\frac{1}{2}} p_i = u_i' C_1 u_i,$$

where the first equality uses the definition $y_t^{(i)}$, and the last equality uses the definition of p_i . Since u_i is an eigenvector of C_1 ,

$$u_i' C_1 u_i = a_i u_i' u_i = a_i,$$

where a_i is the i th largest (in magnitude) eigenvalue of C_1 . Moreover,

$$\mathbb{E} \left[y_t^{(i)2} \right] = p_i' \Gamma_0 p_i = u_i' u_i = 1.$$

Therefore,

$$\rho_i \equiv \mathbb{E} \left[y_t^{(i)} y_{t-1}^{(i)} \right] / \sqrt{\mathbb{E} \left[y_t^{(i)2} \right]} = a_i.$$

The proposition follows the fact that a_i is the i th largest eigenvalue of C_1 in magnitude. \square

Proof of Propositions 3 and 4

The proofs of the two propositions are almost identical, so I only provide a detailed proof of Proposition 4. The only difference between the two is that Proposition 3 relies on Theorems 5 and 8 instead of Theorems 7 and 9, respectively.

I start by taking v to be an arbitrary n -dimensional vector and computing the autocovariances of $v' y_t$ under the pseudo-true and true models. Define $w \equiv \Gamma_0^{\frac{1}{2}} v$. Under the agent's pseudo-true model,

$$\begin{aligned} E^{d*} [v' y_t v' y_{t-l}] &= v' E^{d*} [y_t y_{t-l}'] v = v' E^{d*} [E_{t-l}^{d*} [y_t] y_{t-l}'] v \\ &= v' \sum_{i=1}^d a_i^l q_i p_i' E^{d*} [y_{t-l} y_{t-l}'] v = \sum_{i=1}^d a_i^l v' q_i p_i' \Gamma_0 v \end{aligned}$$

$$= \sum_{i=1}^d a_i^l v' \Gamma_0^{\frac{1}{2}} u_i u_i' \Gamma_0^{\frac{1}{2}} v = \sum_{i=1}^d a_i^l w' u_i u_i' w,$$

where the first equality is by Theorem 2, the second equality follows the fact that the agent's subjective model satisfies the law of iterated expectations, the third and fifth equalities are by Theorem 7, and the fourth equality is by Theorem 9. On the other hand, under the true model,

$$\mathbb{E}[v' y_t v' y_{t-l}] = v' \mathbb{E}[y_t y_{t-l}'] v = v' \Gamma_l v = v' \left(\frac{\Gamma_l + \Gamma_l'}{2} \right) v = v' \Gamma_0^{\frac{1}{2}} C_l \Gamma_0^{\frac{1}{2}} v = w' C_l w.$$

To prove part (a), set $v = p_1$, which implies $v' y_t = y_t^{(1)}$ and $w = \Gamma_0^{\frac{1}{2}} p_1 = u_1$. Therefore,

$$\left| E^{d*} [y_t^{(1)} y_{t-l}^{(1)}] \right| = \left| \sum_{i=1}^d a_i^l u_1' u_i u_i' u_1 \right| = |a_1^l| = |a_1|^l.$$

Furthermore,

$$\left| \mathbb{E} [y_t^{(1)} y_{t-l}^{(1)}] \right| = |u_1' C_l u_1| \leq \rho(C_l) u_1' u_1 = \rho(C_l) \leq \rho(C_1)^l = |a_1|^l,$$

where the second inequality is using the assumption that the true process is exponentially ergodic, and the last equality is due to the fact that a_1 is the eigenvalue of C_1 largest in magnitude. On the other hand, by Theorem 9, the variance of $y_t^{(1)}$ is the same under the true and pseudo-true m.i.o. d -state models. Therefore, the agent overestimates the magnitude of $y_t^{(1)}$'s autocorrelation at all lags.

In part (b), set $v = p_n$, which implies $v' y_t = y_t^{(n)}$ and $w = \Gamma_0^{\frac{1}{2}} p_n = u_n$. Thus,

$$\left| E^{d*} [y_t^{(n)} y_{t-l}^{(n)}] \right| = \left| \sum_{i=1}^d a_i^l u_n' u_i u_i' u_n \right| = 0,$$

where I am using the fact that $\{u_i\}_{i=1}^n$ is an orthonormal basis and the assumption that $d < n$. Hence, the agent underestimates the magnitude of $y_t^{(n)}$'s autocorrelation at all lags, regardless of the true autocorrelation of $y_t^{(n)}$. \square

Proof of Propositions 5 and 6

The proofs of the two propositions are essentially identical, so I only provide a detailed proof of Proposition 5. By Theorem 7, the forecasts of an agent who uses pseudo-true m.i.o. d -state models are given by

$$E_t^{d*} [y_{t+s}] = \sum_{i=1}^d a_i^s q_i p_i' y_t, \quad (\text{C.52})$$

where a_i is the i th largest eigenvalue of C_1 , u_i denotes the corresponding eigenvector, normalized to have unit norm, $p_i \equiv \Gamma_0^{-\frac{1}{2}} u_i$, and $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$. Since the eigenvalues of C_1 are all distinct, the corresponding eigenvectors are unique (up to a multiplicative constant). Therefore, all agents use the same values of $\{(a_i, p_i, q_i)\}_i$ to forecast.

Consider an arbitrary decision j , made by an agent who uses models of dimension $d(j)$. The agent's optimal action given her pseudo-true m.i.o. d -state model is given by

$$\begin{aligned} x_{jt} &= E_t^{d(j)*} \left[\sum_{s=1}^{\infty} c'_{js} y_{t+s} \right] = \sum_{s=1}^{\infty} c'_{js} E_t^{d(j)*} [y_{t+s}] \\ &= \sum_{s=1}^{\infty} c'_{js} \sum_{i=1}^{d(j)} a_i^s q_i p_i' y_t = \sum_{i=1}^{d(j)} g_{ji} y_t^{(i)}, \end{aligned}$$

where $y_t^{(i)} \equiv p_i' y_t$ as before, $g_{ji} \equiv \sum_{s=1}^{\infty} a_i^s c'_{js} q_i$ is a constant, which is a finite since $\{c_{js}\}_s$ is absolutely summable. Using vector notation, $x_t \equiv (x_{1t}, \dots, x_{|J|t})' \in \mathbb{R}^{|J|}$, I can write the above expression as

$$x_t = G y_t^{(1:D)},$$

where

$$G \equiv \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{|J|} \end{pmatrix} \in \mathbb{R}^{|J| \times D},$$

$$g_j \equiv (g_{j1} \quad g_{j2} \quad \dots \quad g_{jd(j)} \quad 0 \quad \dots \quad 0) \in \mathbb{R}^{1 \times D},$$

and

$$y_t^{(1:D)} \equiv \begin{pmatrix} y_t^{(1)} \\ y_t^{(2)} \\ \vdots \\ y_t^{(D)} \end{pmatrix} \in \mathbb{R}^D.$$

Since $G \in \mathbb{R}^{|J| \times D}$ and $D \leq |J|$, there exists a D -dimensional subspace \mathcal{V} of $\mathbb{R}^{|J|}$ such that the range of G is contained in \mathcal{V} . That x_t is in the range of G then establishes that $x_t \in \mathcal{V}$ and completes the proof. \square

Proof of Proposition 7

I first show that, in a linear equilibrium, r_t^n and μ_t can be written as linear functions of \hat{x}_t , $\hat{\pi}_t$, and \hat{i}_t . Suppose r_t^n and μ_t can be written as linear functions of \hat{x}_t , $\hat{\pi}_t$, and \hat{i}_t . Then by the linear-invariance result, agents' forecasts are the same whether they observe vector $f_t \equiv (\hat{x}_t, \hat{\pi}_t, \hat{i}_t)'$ or vector y_t , consisting of all the observables. Furthermore, since shocks follow an exponentially ergodic process and f_t is an invertible linear transformation of the vector of shocks, f_t follows an exponentially ergodic process as well. Therefore, by the linear-invariance result and Theorem 5,

$$E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = \gamma_x \hat{z}_t, \quad (\text{C.53})$$

$$E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = \gamma_{\pi} \hat{z}_t, \quad (\text{C.54})$$

where γ_x and γ_π are constants that are to be determined in equilibrium, $\hat{z}_t = p' f_t$ is the agents' time- t estimate of the subjective state, and $p \equiv (p_x, p_\pi, p_i)'$ is the relative attention vector. Substituting in (11) and (12) and collecting terms, I get

$$\sigma r_t^n = \hat{x}_t + \sigma \hat{i}_t - \gamma_x (p_x \hat{x}_t + p_\pi \hat{\pi}_t + p_i \hat{i}_t), \quad (\text{C.55})$$

$$\mu_t = \hat{\pi}_t - \kappa \hat{x}_t - \gamma_\pi (p_x \hat{x}_t + p_\pi \hat{\pi}_t + p_i \hat{i}_t). \quad (\text{C.56})$$

These expressions verify my guess that r_t^n and μ_t can be written as linear functions of \hat{x}_t , $\hat{\pi}_t$, and \hat{i}_t .

I next find constants γ_x and γ_π . Using the linear-invariance result to substitute for σr_{t+s}^n and μ_{t+s} from the above equations in (C.53) and (C.54) and using Theorem 5 to characterize the resulting subjective expectations, I get

$$E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = \frac{a((1-\beta\gamma_x p_x)q_x - (\sigma + \beta\gamma_x p_\pi)q_\pi - \beta\gamma_x p_i q_i)}{1-a\beta} \hat{z}_t,$$

$$E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = \frac{a\beta(-\delta\gamma_\pi p_x q_x + (1-\delta\gamma_\pi p_\pi)q_\pi - \delta\gamma_\pi p_i q_i)}{1-a\beta\delta} \hat{z}_t,$$

where a is the perceived persistence, and $q \equiv (q_x, q_\pi, q_i)'$ is the relative sensitivity vector. The above equations give two linear equations for the two unknowns γ_x and γ_π . The solution is given by

$$\gamma_x = a(q_x - \sigma q_\pi),$$

$$\gamma_\pi = a\beta q_\pi,$$

where I am using the fact that $p'q = 1$. Finally, solving equations (C.55) and (C.56) for \hat{x}_t and $\hat{\pi}_t$ results in equations (13) and (14). \square

Proof of Proposition 8

I guess and verify that, in any linear equilibrium, \hat{x}_t and $\hat{\pi}_t$ can be written as linear functions of r_t^n . Since μ_t is identically zero, \hat{i}_t is always equal to r_t^n , and \hat{x}_t and $\hat{\pi}_t$ are linear functions of r_t^n , the vector of observables is a linear function of r_t^n . Furthermore, r_t^n follows an exponentially ergodic process. Therefore, by the linear-invariance result and Theorem 5,

$$E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = \iota_x r_t^n, \quad (\text{C.57})$$

$$E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = \iota_\pi r_t^n, \quad (\text{C.58})$$

for some constants ι_x and ι_π , which are to be determined in equilibrium. Substituting in (11) and (12) and collecting terms, I get

$$\hat{x}_t = \iota_x r_t^n, \quad (\text{C.59})$$

$$\hat{\pi}_t = \kappa \hat{x}_t + \iota_\pi r_t^n = (\kappa \iota_x + \iota_\pi) r_t^n. \quad (\text{C.60})$$

These expressions verify the guess that \hat{x}_t and $\hat{\pi}_t$ are linear functions of r_t^n .

I next find ι_x and ι_π . Using the linear-invariance result to substitute for \hat{x}_{t+s} and $\hat{\pi}_{t+s}$ from the above equations in (C.57) and (C.58) and using Theorem 5 to characterize the resulting subjective expectations, I get

$$\begin{aligned} \iota_x r_t^n &= E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = \frac{a((1-\beta)\iota_x - \sigma(\kappa\iota_x + \iota_\pi))}{1-a\beta} r_t^n, \\ \iota_\pi r_t^n &= E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = \frac{a\beta(\kappa\iota_x + (1-\delta)\iota_\pi)}{1-a\beta\delta} r_t^n, \end{aligned}$$

where a is the perceived persistence. The above equations give two linear equations for the two unknowns ι_x and ι_π , with the unique solution given by $\iota_x = \iota_\pi = 0$. Therefore, in the unique linear equilibrium, the output gap and inflation rate are both identically zero. \square

D Omitted Details for the NK Application

D.1 Forward Guidance

By the linear-invariance result, agents' expectations respect any intratemporal linear relationships that hold in the equilibrium without forward guidance. Therefore, substituting from equations (C.55) and (C.56) in (11) and (12), I get

$$E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \nu'_x f_{t+s} \right], \quad (\text{D.1})$$

$$E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \nu'_\pi f_{t+s} \right], \quad (\text{D.2})$$

where $\nu_x, \nu_\pi \in \mathbb{R}^3$ are vectors that satisfy

$$\begin{aligned} \nu'_x f_t &= \frac{1}{\beta} \left[(1 - \beta\gamma_x p_x) \hat{x}_t - (\sigma + \beta\gamma_x p_\pi) \hat{\pi}_t - \beta\gamma_x p_i \hat{i}_t \right], \\ \nu'_\pi f_t &= \frac{1}{\delta} \left[-\delta\gamma_\pi p_x \hat{x}_t + (1 - \delta\gamma_\pi p_\pi) \hat{\pi}_t - \delta\gamma_\pi p_i \hat{i}_t \right]. \end{aligned}$$

On the other hand,

$$E_t^{1*} [f_{t+s}] = \Sigma_{f_s \omega_T} \Sigma_{\omega_T \omega_T}^{-1} \omega_T, \quad (\text{D.3})$$

where $\omega_T \equiv (f'_t, \hat{i}_{t+1}, \dots, \hat{i}_{t+T})' \in \mathbb{R}^{3+T}$, $\Sigma_{f_s \omega_T} \equiv E^{1*} [f_{t+s} \omega'_T]$, and $\Sigma_{\omega_T \omega_T} \equiv E^{1*} [\omega_T \omega'_T]$. Therefore,

$$\begin{aligned} E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \nu'_x f_{t+s} \right] &= \psi'_{xT} \omega_T, \\ E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \nu'_\pi f_{t+s} \right] &= \psi'_{\pi T} \omega_T. \end{aligned}$$

where $\psi_{xT}, \psi_{\pi T} \in \mathbb{R}^{3+T}$ are vectors defined as

$$\psi'_{xT} \equiv (\psi'_{xf}, \psi_{xi_1}, \dots, \psi_{xi_T})' \equiv \nu'_x \left(\sum_{s=1}^{\infty} \beta^s \Sigma_{f_s} \omega_T \right) \Sigma_{\omega_T \omega_T}^{-1}, \quad (\text{D.4})$$

$$\psi'_{\pi T} \equiv (\psi'_{\pi f}, \psi_{\pi i_1}, \dots, \psi_{\pi i_T})' \equiv \nu'_{\pi} \left(\sum_{s=1}^{\infty} (\beta \delta)^s \Sigma_{f_s} \omega_T \right) \Sigma_{\omega_T \omega_T}^{-1}, \quad (\text{D.5})$$

and $\psi_{xf} \equiv (\psi_{xx}, \psi_{x\pi}, \psi_{xi})'$ and $\psi_{\pi f} \equiv (\psi_{\pi x}, \psi_{\pi\pi}, \psi_{\pi i})'$ are vectors in \mathbb{R}^3 . Therefore,

$$\begin{aligned} E_t^{1*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma (\hat{i}_{t+s} - r_{t+s}^n) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] &= \psi'_{xf} f_t + \sum_{s=1}^T \psi_{xi_s} \hat{i}_{t+s}, \\ E_t^{1*} \left[\sum_{s=1}^{\infty} (\beta \delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] &= \psi'_{\pi f} f_t + \sum_{s=1}^T \psi_{\pi i_s} \hat{i}_{t+s}. \end{aligned}$$

Substituting in equations (11) and (12), I get

$$\begin{aligned} \hat{x}_t &= -\sigma (\hat{i}_t - r_t^n) + \psi_{xx} \hat{x}_t + \psi_{x\pi} \hat{\pi}_t + \psi_{xi} \hat{i}_t + \sum_{s=1}^T \psi_{xi_s} \hat{i}_{t+s}, \\ \hat{\pi}_t &= \kappa \hat{x}_t + \mu_t + \psi_{\pi x} \hat{x}_t + \psi_{\pi\pi} \hat{\pi}_t + \psi_{\pi i} \hat{i}_t + \sum_{s=1}^T \psi_{\pi i_s} \hat{i}_{t+s}. \end{aligned}$$

The above equations can be solved for \hat{x}_t and \hat{i}_t to get,

$$\begin{aligned} \hat{x}_t &= \nu_{xi} \hat{i}_t + \nu_{xn} r_t^n + \nu_{x\mu} \mu_t + \sum_{s=1}^T \nu_{xi_s} \hat{i}_{t+s}, \\ \hat{\pi}_t &= \nu_{\pi i} \hat{i}_t + \nu_{\pi n} r_t^n + \nu_{\pi\mu} \mu_t + \sum_{s=1}^T \nu_{\pi i_s} \hat{i}_{t+s}, \end{aligned}$$

for some constants that depend on the ψ 's.

It only remains to compute ψ_{xT} and $\psi_{\pi T}$. I first compute the elements of $\Sigma_{f_s \omega_T}$. By the law of iterated expectations and Theorem 5,

$$E^{1*}[f_{t+s} f_t'] = E^{1*}[E^{1*}[f_{t+s}|f_t] f_t'] = E^{1*}[a^s q p' f_t f_t'] = a^s q p' \Gamma_0.$$

Next consider elements of the form $E^{1*}[f_{t+s} \hat{i}_{t+\tau}]$. If $s < \tau$, then

$$E^{1*}[f_{t+s} \hat{i}_{t+\tau}] = E^{1*}[f_{t+s} E^{1*}[\hat{i}_{t+\tau}|f_{t+s}]] = E^{1*}[f_{t+s} a^{\tau-s} q_i p' f_{t+s}] = a^{\tau-s} q_i E^{1*}[f_{t+s} f_{t+s}'] p = a^{\tau-s} q_i \Gamma_0 p.$$

Likewise, if $s > \tau$, then

$$E^{1*}[f_{t+s} \hat{i}_{t+\tau}] = E^{1*}[\hat{i}_{t+\tau} E^{1*}[f_{t+s}|f_{t+\tau}]] = E^{1*}[e_i' f_{t+\tau} a^{s-\tau} q p' f_{t+\tau}] = a^{s-\tau} q p' E^{1*}[f_{t+s} f_{t+s}'] e_i = a^{s-\tau} q p' \Gamma_0 e_i,$$

where e_i is the coordinate vector that selects element \hat{i}_t of vector $f_t = (\hat{o}_t, \hat{\pi}_t, \hat{i}_t)$, i.e., $\hat{i}_t = e_i' f_t$.

Finally, if $s = \tau$, then

$$E^{1*}[f_{t+s} \hat{i}_{t+\tau}] = E^{1*}[f_{t+s} f_{t+s}' e_i] = \Gamma_0 e_i.$$

I next compute the elements of $\Sigma_{\omega_T \omega_T}$. First, note that

$$E^{1*}[f_t f_t'] = \Gamma_0,$$

and

$$E^{1*}[f_t' \hat{i}_{t+\tau}] = E^{1*}[f_t' E^{1*}[\hat{i}_{t+\tau}|f_t]] = E^{1*}[a^\tau q_i p' f_t f_t'] = a^\tau q_i p' \Gamma_0.$$

Finally, if $\tau < \tau'$, then

$$E^{1*}[\hat{i}_{t+\tau} \hat{i}_{t+\tau'}] = E^{1*}[\hat{i}_{t+\tau} E^{1*}[\hat{i}_{t+\tau'}|f_{t+\tau}]] = E^{1*}[e_i' f_{t+\tau} a^{\tau'-\tau} q_i p' f_{t+\tau}] = a^{\tau'-\tau} q_i p' \Gamma_0 e_i,$$

and

$$E^{1*}[\hat{i}_{t+\tau} \hat{i}_{t+\tau}] = e_i' E^{1*}[f_{t+\tau} f_{t+\tau}'] e_i = e_i' \Gamma_0 e_i.$$

Putting everything together, I get

$$\Sigma_{\omega_T \omega_T} = \begin{pmatrix} \Gamma_0 & a q_i \Gamma_0 p & a^2 q_i \Gamma_0 p & \dots & a^T q_i \Gamma_0 p \\ a q_i p' \Gamma_0 & e_i' \Gamma_0 e_i & a q_i p' \Gamma_0 e_i & \dots & a^{T-1} q_i p' \Gamma_0 e_i \\ a^2 q_i p' \Gamma_0 & a q_i p' \Gamma_0 e_i & e_i' \Gamma_0 e_i & \dots & a^{T-2} q_i p' \Gamma_0 e_i \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^T q_i p' \Gamma_0 & a^{T-1} q_i p' \Gamma_0 e_i & a^{T-2} q_i p' \Gamma_0 e_i & \dots & e_i' \Gamma_0 e_i \end{pmatrix}. \quad (\text{D.6})$$

and

$$\Sigma_{f_s \omega_T} = \begin{cases} (a^s q p' \Gamma_0 & \Gamma_0 e_i & a q_i \Gamma_0 p & a^2 q_i \Gamma_0 p & \dots & a^{T-1} q_i \Gamma_0 p) & \text{if } s = 1, \\ (a^s q p' \Gamma_0 & a^{s-1} q p' \Gamma_0 e_i & \dots & a q p' \Gamma_0 e_i & \Gamma_0 e_i & a q_i \Gamma_0 p & \dots & a^{T-s} q_i \Gamma_0 p) & \text{if } 1 < s < T, \\ (a^s q p' \Gamma_0 & a^{s-1} q p' \Gamma_0 e_i & \dots & a q p' \Gamma_0 e_i & \Gamma_0 e_i) & & & \text{if } s = T, \\ (a^s q p' \Gamma_0 & a^{s-1} q p' \Gamma_0 e_i & \dots & a^{s-T} q p' \Gamma_0 e_i) & & & & \text{if } s > T. \end{cases} \quad (\text{D.7})$$

Therefore,

$$\begin{aligned} & \sum_{s=1}^{\infty} \beta^s \Sigma_{f_s \omega_T} \\ &= \left(\sum_{s=1}^{\infty} (a\beta)^s q p' \Gamma_0 \quad \beta \Gamma_0 e_i + \sum_{s=2}^{\infty} a^{s-1} \beta^s q p' \Gamma_0 e_i \quad \dots \quad \sum_{s=1}^{T-1} a^{T-s} \beta^s q_i \Gamma_0 p + \beta^T \Gamma_0 e_i + \sum_{s=T+1}^{\infty} a^{s-T} \beta^s q p' \Gamma_0 e_i \right) \\ &= \left(\frac{a\beta q p' \Gamma_0}{1-a\beta} \quad \beta \Gamma_0 e_i + \frac{a\beta^2 q p' \Gamma_0 e_i}{1-a\beta} \quad a\beta q_i \Gamma_0 p + \beta^2 \Gamma_0 e_i + \frac{a\beta^3 q p' \Gamma_0 e_i}{1-a\beta} \quad \dots \quad \frac{(a^T \beta - \beta^T a) q_i \Gamma_0 p}{a-\beta} + \beta^T \Gamma_0 e_i + \frac{a\beta^{T+1} q p' \Gamma_0 e_i}{1-a\beta} \right). \end{aligned}$$

Likewise,

$$\sum_{s=1}^{\infty} (\beta\delta)^s \Sigma_{f_s \omega_T} = \left(\frac{a\beta\delta q p' \Gamma_0}{1-a\beta\delta} \quad \beta\delta \Gamma_0 e_i + \frac{a(\beta\delta)^2 q p' \Gamma_0 e_i}{1-a\beta\delta} \quad \dots \quad \frac{(a^T \beta\delta - (\beta\delta)^T a) q_i \Gamma_0 p}{a-\beta\delta} + (\beta\delta)^T \Gamma_0 e_i + \frac{a(\beta\delta)^{T+1} q p' \Gamma_0 e_i}{1-a\beta\delta} \right).$$

Given the expressions for $\Sigma_{\omega_T \omega_T}$, $\sum_{s=1}^{\infty} \beta^s \Sigma_{f_s \omega_T}$, and $\sum_{s=1}^{\infty} (\beta\delta)^s \Sigma_{f_s \omega_T}$, one can use (D.4) and (D.5) to find ψ_{xT} and $\psi_{\pi T}$.

D.2 Estimation

I choose the variance-covariance and lag-one autocovariance of $s_t \equiv (\hat{i}_t, r_t^n, \mu_t)'$ to match the variance-covariance and lag-one autocovariance of $f_t = (\hat{x}_t, \hat{\pi}_t, \hat{i}_t)'$. The estimated values are given by

$$E[s_t s_t'] = \begin{pmatrix} 10.9 & 16.4 & 0.200 \\ 16.4 & 32.1 & -0.0827 \\ 0.200 & -0.0827 & 0.0994 \end{pmatrix},$$

and

$$E[s_t s_{t-1}'] = \begin{pmatrix} 10.4 & 16.2 & 0.155 \\ 15.0 & 30.7 & -0.146 \\ 0.302 & 0.129 & 0.0920 \end{pmatrix}.$$

Figure D.1 plots $\rho(C_l)$ in solid red and $\rho(C_1)^l$ in dashed green, where $\rho(C_l)$ denotes the spectral radius of the lag- l autocorrelation matrix of f_t .⁵¹ The figure verifies that the estimated process is exponentially ergodic, and so, the agents' pseudo-true model is described by Theorem 5.

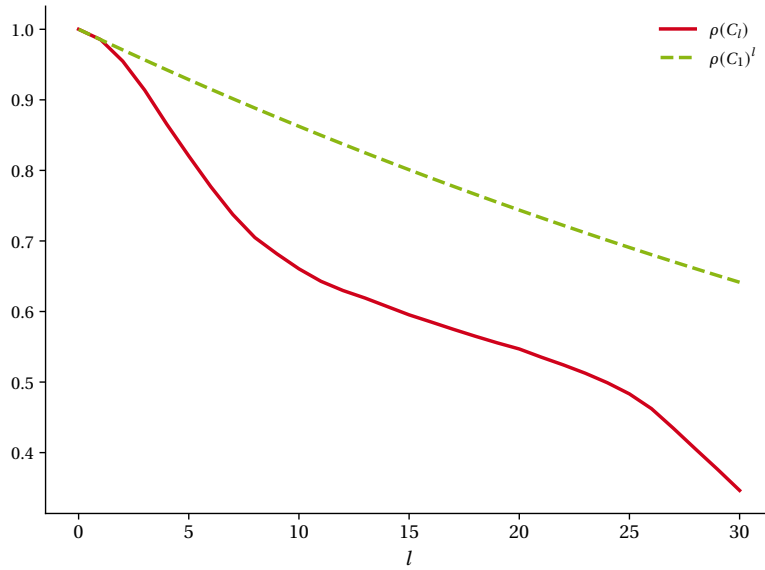


Figure D.1. Test of Exponential Ergodicity

E Omitted Details for the RBC Application

E.1 Temporary Equilibrium

The (log-)linearized temporary-equilibrium conditions are given by

$$\hat{o}_t = \hat{a}_t + \alpha \hat{k}_t + (1 - \alpha) \hat{n}_t, \quad (\text{E.1})$$

$$\hat{w}_t = \hat{a}_t + \alpha (\hat{k}_t - \hat{n}_t), \quad (\text{E.2})$$

⁵¹The result would be identical if I instead used the autocorrelation matrices of s_t . This is a corollary of the linear-invariance result.

$$\hat{r}_t = r\hat{a}_t + (1 - \alpha)r(\hat{n}_t - \hat{k}_t), \quad (\text{E.3})$$

$$\hat{n}_t = \frac{1}{\varphi}\hat{w}_t - \frac{1}{\sigma\varphi}\hat{c}_t, \quad (\text{E.4})$$

$$\hat{i}_t = \frac{o}{i}\hat{y}_t - \frac{c}{i}\hat{c}_t, \quad (\text{E.5})$$

$$\hat{k}_t = (1 - \delta)\hat{k}_{t-1} + \delta\hat{i}_{t-1}, \quad (\text{E.6})$$

$$\hat{a}_t = \rho\hat{a}_{t-1} + \epsilon_t, \quad (\text{E.7})$$

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - \sigma\beta E_t[\hat{r}_{t+1}], \quad (\text{E.8})$$

where \hat{r}_t denotes the first-order deviation of the rental rate of capital from its steady-state value and the remaining hatted variables are log-deviations from the corresponding steady-state values. The Euler equation (E.8) may not hold away from rational expectations if \hat{c}_t denotes the aggregate consumption; it is valid under arbitrary expectations only if \hat{c}_t denotes individual consumption. However, the individual consumption Euler equation can be combined with the households' intertemporal budget constraint and the transversality condition to obtain an aggregate consumption function that is valid under arbitrary expectations. The log-linearized household budget constraint is given by

$$\hat{k}_{t+1} = (1 - \delta + r)\hat{k}_t + \hat{r}_t + \frac{r(1 - \alpha)}{\alpha}(\hat{w}_t + \hat{n}_t) - \frac{c}{k}\hat{c}_t.$$

Substituting for labor supply in the budget constraint, I get

$$\hat{k}_{t+1} = \frac{1}{\beta}\hat{k}_t + \hat{r}_t + \frac{(1 - \alpha)(1 + \varphi)r}{\alpha\varphi}\hat{w}_t - \left(\frac{(1 - \alpha)r}{\alpha\sigma\varphi} + \frac{c}{k}\right)\hat{c}_t,$$

where I am using the fact that $1 - \delta + r = \beta^{-1}$. Multiplying the above equation by β^t , summing over t , and taking subjective expectations of both sides, I get

$$\left(\frac{(1 - \alpha)r}{\alpha\sigma\varphi} + \frac{c}{k}\right) \sum_{s=0}^{\infty} \beta^s E_t[\hat{c}_{t+s}] = \frac{1}{\beta}\hat{k}_t + \sum_{s=0}^{\infty} \beta^s \left(E_t[\hat{r}_{t+s}] + \frac{(1 - \alpha)(1 + \varphi)r}{\alpha\varphi} E_t[\hat{w}_{t+s}]\right).$$

Define

$$\chi \equiv (1 - \beta) \left(\frac{(1 - \alpha)r}{\alpha\sigma\varphi} + \frac{c}{k}\right)^{-1},$$

$$\zeta \equiv \frac{(1 - \alpha)(1 + \varphi)r}{\alpha\varphi}.$$

Then the above equation can be written as

$$\frac{1 - \beta}{\chi} \sum_{s=0}^{\infty} \beta^s E_t[\hat{c}_{t+s}] = \frac{1}{\beta}\hat{k}_t + \sum_{s=0}^{\infty} \beta^s E_t[\hat{r}_{t+s}] + \zeta \sum_{s=0}^{\infty} \beta^s E_t[\hat{w}_{t+s}]. \quad (\text{E.9})$$

On the other hand, the Euler equation implies

$$E_t[\hat{c}_{t+s}] = \hat{c}_t + \sigma\beta \sum_{\tau=1}^s E_t[\hat{r}_{t+\tau}].$$

Therefore,

$$\begin{aligned}
\sum_{s=0}^{\infty} \beta^s E_t[\hat{c}_{t+s}] &= \sum_{s=0}^{\infty} \beta^s \hat{c}_t + \sigma\beta \sum_{s=1}^{\infty} \sum_{\tau=1}^s \beta^s E_t[\hat{r}_{t+s}] \\
&= \frac{1}{1-\beta} \hat{c}_t + \sigma\beta \sum_{\tau=1}^{\infty} \sum_{s=\tau}^{\infty} \beta^s E_t[\hat{r}_{t+\tau}] \\
&= \frac{1}{1-\beta} \hat{c}_t + \frac{\beta\sigma}{1-\beta} \sum_{\tau=1}^{\infty} \beta^{\tau} E_t[\hat{r}_{t+\tau}].
\end{aligned}$$

Combining the above with equation (E.9), I get

$$\hat{c}_t = \frac{\chi}{\beta} \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + (\chi - \beta\sigma) \sum_{s=1}^{\infty} \beta^s E_t[\hat{r}_{t+s}] + \chi \zeta \sum_{s=1}^{\infty} \beta^s E_t[\hat{w}_{t+s}]. \quad (\text{E.10})$$

E.2 Constrained Rational Expectations Equilibrium

Suppose households use a pseudo-true one-state model to forecast the wage rate and the rental rate of capital. Define $\omega_t \equiv (\hat{o}_t, \hat{n}_t, \hat{w}_t, \hat{r}_t, \hat{c}_t, \hat{l}_t)$, $f_t \equiv (\hat{k}_t, \hat{a}_t)'$, and $\xi_t \equiv (f_t', \omega_t')'$. Let $v \in \mathbb{R}^8$ be a vector that satisfies

$$v' \xi_t = (\chi - \beta\sigma) \hat{r}_t + \chi \zeta \hat{w}_t.$$

Then equation (E.10) can be written as

$$\hat{c}_t = \frac{\chi}{\beta} \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + \sum_{s=1}^{\infty} \beta^s v' E_t[\xi_{t+s}].$$

Suppose $\xi_t = T f_t$ for some full-rank matrix T —I later verify that this is indeed the case. Then by the linear-invariance result,

$$\hat{c}_t = \frac{\chi}{\beta} \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + \sum_{s=1}^{\infty} \beta^s v' T E_t[f_{t+s}].$$

The households' forecasts of f_t when they use model θ is given by (4). This can be written recursively as

$$E_t^{\theta}[f_{t+s}] = a^s (1 - \eta) \hat{z}_t q, \quad (\text{E.11})$$

$$\hat{z}_t = a\eta \hat{z}_{t-1} + p' f_t = a\eta \hat{z}_{t-1} + p_k \hat{k}_t + p_a \hat{a}_t, \quad (\text{E.12})$$

where \hat{z}_t denote the households' estimate of the subjective state at time t . Therefore,

$$\hat{c}_t = \frac{\chi}{\beta} \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + \frac{a\beta(1-\eta)}{1-a\beta} v' T q \hat{z}_t. \quad (\text{E.13})$$

I guess that $\eta = 0$ in equilibrium and later verify this guess. Solving for \hat{z}_t from (E.12) and substituting in (E.13), I get

$$\hat{c}_t = \left(\frac{\chi}{\beta} + \gamma_k \right) \hat{k}_t + \chi \hat{r}_t + \chi \zeta \hat{w}_t + \gamma_a \hat{a}_t, \quad (\text{E.14})$$

where

$$\gamma_k \equiv \frac{a\beta}{1-a\beta} \nu' T q p_k, \quad (\text{E.15})$$

$$\gamma_a \equiv \frac{a\beta}{1-a\beta} \nu' T q p_a. \quad (\text{E.16})$$

Equations (E.1)–(E.5) and (E.14) can be solved for ω_t as a function of f_t . This verifies the guess that $\xi_t = (f'_t, \omega'_t)' = T f_t$ and leads to an expression for matrix T . In particular,

$$\hat{l}_t = \psi_k \hat{k}_t + \psi_a \hat{a}_t,$$

for some ψ_k and ψ_a . Substituting for \hat{l}_{t-1} from above in (E.6), I get

$$\hat{k}_t = (1 - \delta + \delta \psi_k) \hat{k}_{t-1} + \delta \psi_a \hat{a}_{t-1}. \quad (\text{E.17})$$

I can now describe the constrained rational expectations equilibrium. Equations (E.7) and (E.17) can be written in vector form as

$$f_t = \mathbb{F}(\gamma_k, \gamma_a) f_{t-1} + \epsilon_t. \quad (\text{E.18})$$

An equilibrium is given by tuples (γ_k^*, γ_a^*) and (a^*, η^*, p^*, q^*) such that (i) (a^*, η^*, p^*, q^*) is the pseudo-true one-state model when the true process is given by (E.18) with $\gamma_k = \gamma_k^*$ and $\gamma_a = \gamma_a^*$, (ii) γ_k^* and γ_a^* are given by equations (E.15) and (E.16) for $a = a^*$, $p = p^*$, and $q = q^*$, and (iii) $\eta^* = 0$.

Finding an equilibrium requires solving a fixed-point equation. I start with a candidate $(\gamma_k, \gamma_a, \eta)$, with $\eta = 0$. The candidate defines a true process as in (E.18). This process in turn leads to a pseudo-true one-state model $(\tilde{a}, \tilde{\eta}, \tilde{p}, \tilde{q})$. Such a pseudo-true one-state model, in turn, defines a $(\tilde{\gamma}_k, \tilde{\gamma}_a)$ pair through equations (E.15) and (E.16). I solve for the equilibrium by numerically minimizing the Euclidean distance between tuples $(\tilde{\gamma}_k, \tilde{\gamma}_a, \tilde{\eta})$ and $(\gamma_k, \gamma_a, \eta)$ over the set of all (γ_k, γ_a) pairs. The fixed-point turns out to satisfy $\tilde{\eta} = \eta = 0$, verifying my earlier conjecture.

F Omitted Details for the DMP Application

F.1 Non-Linear Equilibrium

I start with the workers' problem. Let U_t and V_t denote the time- t value to a worker of unemployment and employment, respectively. Those random variables solve the following Bellman equations:

$$U_t = b + \beta E_t [p_t V_{t+1} + (1 - p_t) U_{t+1}], \quad (\text{F.1})$$

$$V_t = w_t + \beta E_t [s_t U_{t+1} + (1 - s_t) V_{t+1}], \quad (\text{F.2})$$

where b denotes the workers' flow payoff from being unemployed, w_t denotes the wage rate, and $p_t = \mu \theta_t^{1-\alpha}$ denotes the job-finding probability, with θ_t the labor market tightness and μ and α

parameters of the matching function. Subtracting U_t from V_t , I get

$$V_t - U_t = w_t - b + \beta E_t [(1 - s_t - p_t)(V_{t+1} - U_{t+1})]. \quad (\text{F.3})$$

Define

$$\lambda_{t,t+\tau}^w \equiv \prod_{k=0}^{\tau-1} (1 - s_{t+k} - p_{t+k}). \quad (\text{F.4})$$

Solving (F.3) forward, I get

$$V_t - U_t = w_t - b + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \lambda_{t,t+\tau}^w (w_{t+\tau} - b) \right]. \quad (\text{F.5})$$

This equation is valid under arbitrary expectations.

I consider the firms next. Let J_t denote the time- t value to a firm of a job. It solves the following Bellman equation:

$$J_t = a_t - w_t + \beta E_t [(1 - s_t)J_{t+1}].$$

Solving the equation forward, I get

$$J_t = a_t - w_t + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \lambda_{t,t+\tau}^f (a_{t+\tau} - w_{t+\tau}) \right], \quad (\text{F.6})$$

where

$$\lambda_{t,t+\tau}^f \equiv \prod_{k=0}^{\tau-1} (1 - s_{t+k}), \quad (\text{F.7})$$

and $\lambda_{t,t}^f \equiv 1$. Free entry by firms implies

$$0 = -k + \beta E_t [q_t J_{t+1}], \quad (\text{F.8})$$

where $q_t = \mu \theta_t^{-\alpha}$ is the probability of filling a vacancy in each period. Substituting for J_t in (F.8) from (F.6), I get

$$\theta_t^{\alpha} = \frac{\mu}{k} E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \lambda_{t+1,t+\tau}^f (a_{t+\tau} - w_{t+\tau}) \right]. \quad (\text{F.9})$$

Equation (F.9) determines tightness as a function of the firms' expectations of wage and labor productivity.

The wage rate is determined by Nash bargaining. Under Nash bargaining,

$$\frac{J_t}{1 - \delta} = \frac{V_t - U_t}{\delta},$$

where δ denotes the workers' bargaining power. Combining the above equation with (F.5) and (F.6) and solving for w_t , I get

$$w_t = \delta a_t + (1 - \delta)b + \delta E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \lambda_{t,t+\tau}^f (a_{t+\tau} - w_{t+\tau}) \right] - (1 - \delta) E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} \lambda_{t,t+\tau}^w (w_{t+\tau} - b) \right]. \quad (\text{F.10})$$

The unemployment rate follows the first-order difference equation

$$u_t = u_{t-1} + s_{t-1}(1 - u_{t-1}) - p_{t-1}u_{t-1}. \quad (\text{F.11})$$

F.2 Steady State

I first consider a steady state in which $a_t = 1 > b$, $w_t = w$, $\theta_t = \theta$, $s_t = s$, and agents have perfect foresight. Equation (F10) implies that in the steady state,

$$\frac{(1 - \delta)(w - b)}{1 - \beta(1 - s - p)} = \frac{\delta(1 - w)}{1 - \beta(1 - s)}.$$

Therefore,

$$w = \frac{\delta(1 - \beta(1 - s - p)) + (1 - \delta)(1 - \beta(1 - s))b}{1 - \beta(1 - s - \delta p)}.$$

Equation (F6) implies that the value of a job to a firm is constant in the steady state:

$$J_t = J = \frac{1}{1 - \beta(1 - s)}(1 - w).$$

Equation (F8) and the definition of q_t imply

$$\frac{\mu}{k\theta^\alpha} = \frac{1}{\beta J}.$$

The steady-state unemployment rate satisfies

$$s \frac{1 - u}{u} = p.$$

F.3 Log-Linear Model

I next log-linearize the model around the steady state. Log-linearizing (F4) and (F7), I get

$$\begin{aligned} \hat{\lambda}_{t,t+\tau}^w &= -\frac{1}{1 - s - p} \sum_{k=0}^{\tau-1} (p\hat{p}_{t+k} + s\hat{s}_{t+k}), \\ \hat{\lambda}_{t,t+\tau}^f &= -\frac{1}{1 - s} \sum_{k=0}^{\tau-1} s\hat{s}_{t+k}. \end{aligned}$$

Log-linearizing $p_t = \mu\theta_t^{1-\alpha}$, I get

$$\hat{p}_t = (1 - \alpha)\hat{\theta}_t.$$

Log-linearizing (F9),

$$\hat{\theta}_t = E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau (1 - s)^{\tau-1} \left(\frac{(1 - b)\mu}{\alpha k \theta^\alpha} \hat{a}_{t+\tau} - \frac{w\mu}{\alpha k \theta^\alpha} \hat{w}_{t+\tau} + \frac{(1 - w)\mu}{\alpha k \theta^\alpha} \hat{\lambda}_{t+1,t+\tau} \right) \right].$$

Substituting for $\mu/(k\theta^\alpha)$, I get

$$\hat{\theta}_t = E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau-1} (1 - s)^{\tau-1} \left(\frac{1 - b}{\alpha J} \hat{a}_{t+\tau} - \frac{w}{\alpha J} \hat{w}_{t+\tau} + \frac{1 - w}{\alpha J} \hat{\lambda}_{t+1,t+\tau} \right) \right].$$

The term involving $\hat{\lambda}_{t,t+\tau}^f$ can be simplified further:

$$\sum_{\tau=1}^{\infty} \beta^{\tau-1} (1 - s)^{\tau-1} \hat{\lambda}_{t+1,t+\tau} = - \sum_{\tau=2}^{\infty} \beta^{\tau-1} (1 - s)^{\tau-1} \frac{s}{1 - s} \sum_{k=0}^{\tau-2} \hat{s}_{t+1+k}$$

$$\begin{aligned}
&= -s \sum_{k=0}^{\infty} \hat{s}_{t+1+k} \sum_{\tau=k+2}^{\infty} \beta^{\tau-1} (1-s)^{\tau-2} \\
&= -\frac{\beta s}{1-\beta(1-s)} \sum_{\tau=1}^{\infty} \beta^{\tau-1} (1-s)^{\tau-1} \hat{s}_{t+\tau}.
\end{aligned}$$

Define

$$\zeta \equiv \frac{\beta s(1-w)}{1-\beta(1-s)}.$$

Then,

$$\hat{\theta}_t = \frac{1}{\alpha J} E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau-1} (1-s)^{\tau-1} ((1-b)\hat{a}_{t+\tau} - w\hat{w}_{t+\tau} - \zeta \hat{s}_{t+\tau}) \right]. \quad (\text{F.12})$$

Log-linearizing (F.10),

$$\begin{aligned}
w\hat{w}_t &= \delta(1-b)\hat{a}_t + \delta E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s)^{\tau} ((1-b)\hat{a}_{t+\tau} - w\hat{w}_{t+\tau} + (1-w)\hat{\lambda}_{t,t+\tau}^f) \right] \\
&\quad - (1-\delta) E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} ((w-b)\hat{\lambda}_{t,t+\tau}^w + w\hat{w}_{t+\tau}) \right].
\end{aligned} \quad (\text{F.13})$$

The terms involving $\hat{\lambda}_{t,t+\tau}^w$ and $\hat{\lambda}_{t,t+\tau}^f$ can be simplified further:

$$\begin{aligned}
\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} \hat{\lambda}_{t,t+\tau}^w &= - \sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} \frac{1}{1-s-p} \sum_{k=0}^{\tau-1} (p\hat{p}_{t+k} + s\hat{s}_{t+k}) \\
&= -\beta \sum_{k=0}^{\infty} (p\hat{p}_{t+k} + s\hat{s}_{t+k}) \sum_{\tau=k+1}^{\infty} (\beta(1-s-p))^{\tau-1} \\
&= -\frac{\beta}{1-\beta(1-s-p)} \sum_{k=0}^{\infty} \beta^k (1-s-p)^k (p\hat{p}_{t+k} + s\hat{s}_{t+k}). \\
\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s)^{\tau} \hat{\lambda}_{t,t+\tau}^f &= -\frac{\beta s}{1-\beta(1-s)} \sum_{k=0}^{\infty} \beta^k (1-s)^k \hat{s}_{t+k}.
\end{aligned}$$

Define

$$\chi \equiv \frac{\beta(1-\delta)(w-b)}{1-\beta(1-s-p)}$$

Then, (F.13) can be written as

$$\begin{aligned}
w\hat{w}_t &= \delta(1-b)\hat{a}_t + (s\chi - \delta\zeta)\hat{s}_t + p\chi(1-\alpha)\hat{\theta}_t \\
&\quad + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s)^{\tau} (\delta(1-b)\hat{a}_{t+\tau} - \delta w\hat{w}_{t+\tau} - \delta\zeta \hat{s}_{t+\tau}) \right] \\
&\quad - E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} ((1-\delta)w\hat{w}_{t+\tau} - p\chi(1-\alpha)\hat{\theta}_{t+\tau} - s\chi \hat{s}_{t+\tau}) \right].
\end{aligned} \quad (\text{F.14})$$

Finally, log-linearizing (F.11),

$$\hat{u}_t = (1-s-p)\hat{u}_{t-1} - (1-\alpha)p\hat{\theta}_{t-1} + p\hat{s}_{t-1}. \quad (\text{F.15})$$

F.4 Rational Expectations Equilibrium

I guess and verify that under rational expectations $\hat{\theta}_t = \gamma_{\theta a} \hat{a}_t + \gamma_{\theta s} \hat{s}_t$ and $w\hat{w}_t = \gamma_{wa} \hat{a}_t + \gamma_{ws} \hat{s}_t$. Substituting in (F.12) and (F.14), I get

$$\hat{\theta}_t = \frac{\rho_a}{1 - \beta\rho_a(1-s)} \frac{1-b-\gamma_{wa}}{\alpha J} \hat{a}_t - \frac{\rho_s}{1 - \beta\rho_s(1-s)} \frac{\zeta + \gamma_{ws}}{\alpha J} \hat{s}_t,$$

and

$$\begin{aligned} w\hat{w}_t = & \left[\delta(1-b) + p\chi(1-\alpha)\gamma_{\theta a} + \frac{\beta\delta\rho_a(1-s)(1-b-\gamma_{wa})}{1 - \beta\rho_a(1-s)} \right] \hat{a}_t \\ & + \left[\frac{\beta\rho_a(1-s-p)}{1 - \beta\rho_a(1-s-p)} (p\chi(1-\alpha)\gamma_{\theta a} - (1-\delta)\gamma_{wa}) \right] \hat{a}_t \\ & + \left[s\chi - \delta\zeta + p\chi(1-\alpha)\gamma_{\theta s} - \frac{\beta\delta\rho_s(1-s)(\zeta + \gamma_{ws})}{1 - \beta\rho_s(1-s)} \right] \hat{s}_t \\ & + \left[\frac{\beta\rho_s(1-s-p)}{1 - \beta\rho_s(1-s-p)} (p\chi(1-\alpha)\gamma_{\theta s} + s\chi - (1-\delta)\gamma_{ws}) \right] \hat{s}_t. \end{aligned}$$

These equations validate the guess and yield four linear equations for the four unknowns $\gamma_{\theta a}$, $\gamma_{\theta s}$, γ_{wa} , and γ_{ws} , which can be solved given values for the exogenous parameters. The rational expectations equilibrium is then described by (25) and (F.15) with $\hat{w}_t = \gamma_{wa}^* \hat{a}_t + \gamma_{ws}^* \hat{s}_t$, $\hat{\theta}_t = \gamma_{\theta a}^* \hat{a}_t + \gamma_{\theta s}^* \hat{s}_t$, and $(\gamma_{\theta a}^*, \gamma_{\theta s}^*, \gamma_{wa}^*, \gamma_{ws}^*)$ the solution to the above linear equations.

F.5 Constrained Rational Expectations Equilibrium

I next consider the equilibrium where agents are constrained to use pseudo-true one-state models. I guess (and later verify) that, in equilibrium,

$$\begin{aligned} \hat{\theta}_t &= \psi_{\theta u} \hat{u}_t + \psi_{\theta a} \hat{a}_t + \psi_{\theta s} \hat{s}_t, \\ w\hat{w}_t &= \psi_{wu} \hat{u}_t + \psi_{wa} \hat{a}_t + \psi_{ws} \hat{s}_t. \end{aligned}$$

Using the linear-invariance result to substitute for $\hat{\theta}_{t+\tau}$ and $\hat{w}_{t+\tau}$ in (F.12) and (F.14), I get

$$\hat{\theta}_t = \frac{1}{\alpha J} E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau-1} (1-s)^{\tau-1} ((1-b-\psi_{wa}) \hat{a}_{t+\tau} - \psi_{wu} \hat{u}_{t+\tau} - (\zeta + \psi_{ws}) \hat{s}_{t+\tau}) \right], \quad (\text{F.16})$$

and

$$\begin{aligned} w\hat{w}_t = & p\chi(1-\alpha)\psi_{\theta u} \hat{u}_t + (\delta(1-b) + p\chi(1-\alpha)\psi_{\theta a}) \hat{a}_t + (s\chi - \delta\zeta + p\chi(1-\alpha)\psi_{\theta s}) \hat{s}_t \\ & + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s)^{\tau} (\delta(1-b-\psi_{wa}) \hat{a}_{t+\tau} - \delta\psi_{wu} \hat{u}_{t+\tau} - \delta(\zeta + \psi_{ws}) \hat{s}_{t+\tau}) \right] \\ & + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} (p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu}) \hat{u}_{t+\tau} \right] \\ & + E_t \left[\sum_{\tau=1}^{\infty} \beta^{\tau} (1-s-p)^{\tau} (p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa}) \hat{a}_{t+\tau} \right] \end{aligned}$$

$$+ E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau (1-s-p)^\tau (p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)\hat{s}_{t+\tau} \right]. \quad (\text{F.17})$$

The agents' forecasts are given by equation (4). I guess that $\eta = 0$ in equilibrium and later verify this guess. Given the guess,

$$\begin{aligned} E_t[\hat{u}_{t+\tau}] &= a^\tau q_u p_u \hat{u}_t + a^\tau q_u p_a \hat{a}_t + a^\tau q_u p_s \hat{s}_t, \\ E_t[\hat{a}_{t+\tau}] &= a^\tau q_a p_u \hat{u}_t + a^\tau q_a p_a \hat{a}_t + a^\tau q_a p_s \hat{s}_t, \\ E_t[\hat{s}_{t+\tau}] &= a^\tau q_s p_u \hat{u}_t + a^\tau q_s p_a \hat{a}_t + a^\tau q_s p_s \hat{s}_t. \end{aligned}$$

Using the linear-invariance result to substitute for $E_t[\hat{u}_{t+\tau}]$, $E_t[\hat{a}_{t+\tau}]$, and $E_t[\hat{s}_{t+\tau}]$ in (F.16) and (F.17) and collecting terms verifies the guess that $\hat{\theta} = \psi_{\theta u}\hat{u}_t + \psi_{\theta a}\hat{a}_t + \psi_{\theta s}\hat{s}_t$ and $\hat{w}_t = \psi_{wu}\hat{u}_t + \psi_{wa}\hat{a}_t + \psi_{ws}\hat{s}_t$ and leads to the following linear equations for $\psi_{\theta u}$, $\psi_{\theta a}$, $\psi_{\theta s}$, ψ_{wu} , ψ_{wa} , and ψ_{ws} :

$$\psi_{\theta u} = \frac{ap_u}{1-a\beta(1-s)} \left(\frac{1-b-\psi_{wa}}{\alpha J} q_a - \frac{\psi_{wu}}{\alpha J} q_u - \frac{\zeta+\psi_{ws}}{\alpha J} q_s \right), \quad (\text{F.18})$$

$$\psi_{\theta a} = \frac{ap_a}{1-a\beta(1-s)} \left(\frac{1-b-\psi_{wa}}{\alpha J} q_a - \frac{\psi_{wu}}{\alpha J} q_u - \frac{\zeta+\psi_{ws}}{\alpha J} q_s \right), \quad (\text{F.19})$$

$$\psi_{\theta s} = \frac{ap_s}{1-a\beta(1-s)} \left(\frac{1-b-\psi_{wa}}{\alpha J} q_a - \frac{\psi_{wu}}{\alpha J} q_u - \frac{\zeta+\psi_{ws}}{\alpha J} q_s \right), \quad (\text{F.20})$$

$$\begin{aligned} \psi_{wu} &= p\chi(1-\alpha)\psi_{\theta u} + \frac{a\beta\delta(1-s)p_u}{1-a\beta(1-s)} [(1-b-\psi_{wa})q_a - \psi_{wu}q_u - (\zeta+\psi_{ws})q_s] \\ &\quad + \frac{a\beta(1-s-p)p_u}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa})q_a + (p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu})q_u] \\ &\quad + \frac{a\beta(1-s-p)p_u}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)q_s], \end{aligned} \quad (\text{F.21})$$

$$\begin{aligned} \psi_{wa} &= \delta(1-b) + p\chi(1-\alpha)\psi_{\theta a} + \frac{a\beta\delta(1-s)p_a}{1-a\beta(1-s)} [(1-b-\psi_{wa})q_a - \psi_{wu}q_u - (\zeta+\psi_{ws})q_s] \\ &\quad + \frac{a\beta(1-s-p)p_a}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa})q_a + (p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu})q_u] \\ &\quad + \frac{a\beta(1-s-p)p_a}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)q_s], \end{aligned} \quad (\text{F.22})$$

$$\begin{aligned} \psi_{ws} &= s\chi - \delta\zeta + p\chi(1-\alpha)\psi_{\theta s} + \frac{a\beta\delta(1-s)p_s}{1-a\beta(1-s)} [(1-b-\psi_{wa})q_a - \psi_{wu}q_u - (\zeta+\psi_{ws})q_s] \\ &\quad + \frac{a\beta(1-s-p)p_s}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa})q_a + (p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu})q_u] \\ &\quad + \frac{a\beta(1-s-p)p_s}{1-a\beta(1-s-p)} [(p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)q_s]. \end{aligned} \quad (\text{F.23})$$

I can now describe the constrained rational expectations equilibrium. Given $\hat{\theta} = \psi_{\theta u}\hat{u}_t + \psi_{\theta a}\hat{a}_t + \psi_{\theta s}\hat{s}_t$, equations (25) and (F.15) can be written in vector form as

$$f_t = \mathbb{F}(\psi_{\theta u}, \psi_{\theta a}, \psi_{\theta s})f_{t-1} + \epsilon_t. \quad (\text{F.24})$$

An equilibrium is then given by tuples $(\psi_{\theta u}^*, \psi_{\theta a}^*, \psi_{\theta s}^*, \psi_{wu}^*, \psi_{wa}^*, \psi_{ws}^*)$ and (a^*, η^*, p^*, q^*) such that (i) (a^*, η^*, p^*, q^*) is the pseudo-true one-state model when the true process is given by (F.24) with

$\psi_{\theta u} = \psi_{\theta u}^*$, $\psi_{\theta a} = \psi_{\theta a}^*$, and $\psi_{\theta s} = \psi_{\theta s}^*$, (ii) $(\psi_{\theta u}^*, \psi_{\theta a}^*, \psi_{\theta s}^*, \psi_{wu}^*, \psi_{wa}^*, \psi_{ws}^*)$ solves (F.18)–(F.23) given $a = a^*$, $p = p^*$, $q = q^*$, and (iii) $\eta^* = 0$.

Finding an equilibrium requires solving a fixed-point equation. I start with a candidate $(\psi_{\theta u}, \psi_{\theta a}, \psi_{\theta s}, \eta)$, with $\eta = 0$. The candidate defines a true process as in (F.24). The process leads to a pseudo-true one-state model $(\tilde{a}, \tilde{\eta}, \tilde{p}, \tilde{q})$. Such a pseudo-true one-state model, in turn, defines a $(\tilde{\psi}_{\theta u}, \tilde{\psi}_{\theta a}, \tilde{\psi}_{\theta s})$ triple through equations (F.18)–(F.23). I solve for the equilibrium by numerically minimizing the Euclidean distance between pairs $(\tilde{\psi}_{\theta u}, \tilde{\psi}_{\theta a}, \tilde{\psi}_{\theta s}, \tilde{\eta})$ and $(\psi_{\theta u}, \psi_{\theta a}, \psi_{\theta s}, \eta)$ over the set of all $(\psi_{\theta u}, \psi_{\theta a}, \psi_{\theta s})$ tuples. The fixed-point turns out to satisfy $\tilde{\eta} = \eta = 0$, verifying my earlier conjecture.