Simple Models and Biased Forecasts[*]

Pooya Molavi[†]

October 28, 2023

This paper proposes a framework in which agents are constrained to use simple 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 intertemporal complexity of a model. Agents are boundedly rational in that they can only consider models that are too simple to nest the true process, yet they use the best model among those considered. I show that using simple models adds persistence to forward-looking decisions and increases the comovement among them. I then explain how this insight can bring the predictions of three workhorse macroeconomic models closer to data. In the new-Keynesian model, forward guidance becomes less powerful. In the real business cycle model, consumption responds more sluggishly to productivity shocks. The Diamond–Mortensen–Pissarides model exhibits more internal propagation and more realistic comovement in response to productivity and separation shocks.

[*]I am grateful to Marios Angeletos, Ian Dew-Becker, Théo Durandard, Laura Gáti, Ali Jadbabaie, Alessandro Pavan, Bruno Strulovici, Alireza Tahbaz-Salehi, Iván Werning, and numerous seminar and conference participants for 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 maintains that agents can forecast the future as if they knew the true model of the economy. The unrealistic nature of the rational-expectations assumption would not be of great concern if predictions of the standard macro models were robust to alternative specifications of expectations. However, the answers to many important questions in macro, ranging from the power of forward guidance to the response of firms' hiring decisions to changes in economic conditions, are sensitive to how agents form their expectations.

This paper studies the forecasting biases resulting from individuals' use of simple models and the macroeconomic implications of those biases. I introduce a framework in which the true data-generating process features complex intertemporal relationships among variables, but agents can only entertain stochastic models with a bound on their intertemporal complexity. Specifically, they only consider stochastic processes that can be represented using a d-dimensional state variable, where d is a parameter that captures the complexity of an agent's model. Agents are boundedly rational; they can only entertain models that are too simple to capture the true process but find the best d-dimensional approximation to it.

The framework has sharp predictions for agents' forecasts and forward-looking actions. I show that agents misperceive *intertemporal* statistical relationships among variables observed over time. In particular, while agents can accurately forecast the most persistent components of those observables, they miss the dynamics of their less persistent components. This bias increases the *persistence* and *comovement* of agents' forward-looking choices. Agents make forecasts that are unresponsive to changes in all but the most persistent components. This unresponsiveness anchors forward-looking actions to the most persistent state variables in the economy, thereby increasing the persistence of those actions. Furthermore, different agents—with different payoffs, facing different decisions, and using models with different dimensions—all agree on the most persistent components. Since forward-looking decisions of such agents are influenced by a limited set of common components, those decisions comove more than they would under rational expectations.

This paper focuses on bounded rationality as a reduction in intertemporal complexity. I make two main assumptions that allow me to focus on this dimension of complexity, abstracting from potential errors agents might make when dealing with other forms of complexity. First, I assume that agents are capable of entertaining *any* stochastic model that has a d-dimensional linear-Gaussian state-space representation. Second, the best model is defined as the stochastic process that minimizes the Kullback–Leibler divergence from the true process. While this deviates from the utility-based notion of model optimality prevalent in the rational-inattention literature, it aligns with the emerging literature on model misspecification in game theory (e.g., Esponda and

Pouzo (2016)). A significant consequence of these simplifying assumptions is a useful linear-invariance property: Expectations formed using simple models respect linear *intratemporal* relationships among observables. Moreover, these assumptions greatly enhance the framework's tractability, rendering it seamlessly applicable even in large-scale macro models, without adding to the computational burden or introducing additional degrees of freedom (beyond d).

Many business-cycle models in macroeconomics struggle to generate empirically plausible degrees of persistence and comovement in endogenous aggregate variables. Two common solutions pursued in the literature are to introduce auxiliary frictions such as habit formation and adjustment costs (e.g., Christiano, Eichenbaum, and Evans (2005)) or rely on shocks without clear empirical counterparts (e.g., Smets and Wouters (2007)). This paper presents a novel and parsimonious solution that relies on agents' bounded rationality. By replacing rational expectations with simple models, this approach narrows the gap between theory and data. It increases persistence and comovement in a way that is universally applicable across applications.

I demonstrate the framework's versatility through its integration into three workhorse models in macroeconomics: the new-Keynesian model, the real business cycle (RBC) model, and the Diamond–Mortensen–Pissarides (DMP) model. In each case, I study how the predictions of the macro model change when agents use simple, one-dimensional models. Besides being three of the most canonical models in macroeconomics, these applications are chosen because they showcase different predictions of the simple models framework.

For the first application, I consider a version of the standard new-Keynesian model in which agents use one-dimensional models. As in the rational-expectations version of the model, the equilibrium has a simple linear representation that can be found in closed form. I use the equilibrium characterization to study the implications of bounded rationality for the conduct of monetary policy. Two main insights arise from the analysis: 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. This occurs because inflation targeting leads to a stable economy with simple intertemporal relationships, in which the constraint on agents' models is not binding. Second, replacing rational expectations with simple models offers a resolution to the so-called forward-guidance puzzle (Del Negro, Giannoni, and Patterson (2023)). Forward guidance is implausibly powerful under rational expectations and rapidly grows more powerful as the duration of guidance increases. Agents who use simple models misperceive the intertemporal statistical relationships among output, inflation, and interest rate. This misperception curtails the power of forward guidance and makes its power largely independent of the duration of guidance.

The second application studies the response of the standard real business cycle model to productivity shocks. When substituting rational expectations with simple models, the persistence of consumption increases in a way that resembles habit formation. To see the intuition for this result, note that consumption is a forward-looking decision in the RBC model. Therefore, it only responds to changes in the most persistent state variable in the economy, the capital stock. Since the capital stock moves sluggishly when there is an increase in productivity, so does consumption. The sluggishness of consumption increases the volatilities of consumption, hours, and investment relative to the rational-expectations benchmark, resulting in an increase in the cost of business cycles.

In the last application, I study how the predictions of the standard search and matching model change when agents use simple models. I consider a standard calibration of the DMP model with labor-productivity and separation-rate shocks in which agents are restricted to using one-dimensional models. Replacing rational expectations with simple models adds persistence to the unemployment rate, number of vacancies, and job-finding rate in response to both shocks, thus improving the internal propagation mechanism of the DMP model. Moreover, it enables the DMP model to generate negative comovement between the unemployment rate and vacancies in response to separation shocks—reversing a counterfactual prediction of the model under rational expectations. This reversal is due to the fact that agents use the single state variable in their model to closely track the evolution of the unemployment rate, the economy's most persistent state variable. Separation-rate shocks increase the unemployment rate, thus making firms pessimistic about the state of the economy and leading them to post fewer vacancies.

Related Literature. This paper belongs to the literature in macroeconomics on deviations from full-information rational expectations (FIRE)—see Woodford (2013) for a survey. The literatures on 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 Wiederholt (2009, 2015), and Gabaix (2014) deviate from the FIRE benchmark by imposing imperfect knowledge of the payoff-relevant variables.[1] 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 (or intertemporal) 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.

A large literature studies the question of whether households, firms, and professional forecasters under- or over-extrapolate from new information. Coibion and Gorodnichenko (2015) provide evidence of under-extrapolation in consensus forecasts for professional forecasters. Bordalo, Gennaioli, Ma, and Shleifer (2020) show that individual forecasts of professional forecasters over-extrapolate from recent news. Broer and Kohlhas (2020) find evidence for both under- and over-extrapolation depending on the aggregate variable being studied. More recently, Angeletos, Huo, and Sastry (2021) find evidence of under-extrapolation at short horizons and over-extrapolation

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

at longer horizons, whereas Afrouzi, Kwon, Landier, Ma, and Thesmar (2021) find evidence of over-extrapolation in a lab setting. In parallel with this empirical literature, many papers have proposed theoretical models of under- and over-extrapolation. 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 over-extrapolate 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 that feature under-extrapolation. The framework proposed in this paper is neither a model of under-extrapolation nor of over-extrapolation; agents who use simple models always under-extrapolate some observables and over-extrapolate others.

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 the Kullback–Leibler 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 since 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.

The state-space models used in this paper are relatives of dynamic-factor models, e.g., Stock and Watson (2011, 2016). However, the two offer two distinct ways of decomposing 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.[2]

Finally, in a follow-up paper, Molavi, Tahbaz-Salehi, and Vedolin (forthcoming) use a closely related framework to study the implications of model misspecification for asset prices and returns.

[2]The 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.

They show that constraining the complexity of investors' models leads to return and forecast-error 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 fit used in the paper. Section 3 contains the paper's characterization results for simple models. Section 4 discusses the implications of using 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. Additional results are provided in three appendices. The proofs of the theoretical results and other calculations are relegated to the online appendices.

2 Framework

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

2.1 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, but the 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 purely non-deterministic, stationary, and ergodic and has a finite second moment. Second, there exists a subspace $\mathcal{W}$ of $\mathbb{R}^n$ (possibly equal to $\mathbb{R}^n$ itself) such that y_t is supported on $\mathcal{W}$ with density $\mathbb{f}$.[3] Finally, the true process has finite entropy rate, i.e., $\lim_{t\to\infty} \frac{1}{t}\mathbb{E}\left[-\log \mathbb{f}(y_1,\dots,y_t)\right] < \infty$. These assumptions are all quite weak. For instance, they are satisfied if y_t follows a stationary vector ARMA process with Gaussian innovations.

The agent has perfect information about the past realizations of the observable; her time-t information set is 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.

[3] This assumption is weaker than the assumption that $\mathbb{P}$ has full support over $\mathbb{R}^n$ because 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.

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

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

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 stochastic process over $\{y_t\}_{t=-\infty}^{\infty}$ 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.[4] I let P^θ denote the d-state model parameterized by the collection of matrices $\theta \equiv (A, B, Q, R)$, let $E^\theta[\cdot]$ denote the corresponding expectation operator, and let Θ_d denote the set of all d-state models.[5] Whenever there is no risk of confusion, I use the term d-state model to refer both to the stochastic process P^θ for y_t and the parameters $\theta = (A, B, Q, R)$ of its state-space representation.

The integer d captures the agent's sophistication in modeling the stochastic process for the vector of observables, with larger values of d indicating agents who can entertain more complex models. When d is sufficiently large, the agent can approximate the unconditional and conditional second moments of any purely non-deterministic covariance-stationary process arbitrarily well using a model in her set of models. 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 Θ_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 the agent's bounded rationality, 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

[4] A 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 stationary ergodic process.

[5] One 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 $E[y_t y'_{t-l}] = CA^{l-1}\overline{C}'$ for all $l = 1, 2, \ldots$, some convergent $d \times d$ matrix A, and some $C, \overline{C} \in \mathbb{R}^{n\times d}$. See, for instance, Faurre (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 agents' view of the state of the economy in the macro applications I consider in this paper.

issues arising from small samples and instead consider the long-run limit where the sampling error vanishes.

2.3 The Notion of Fit

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.[6] 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 subspace $\mathcal{W}$ of $\mathbb{R}^n$. If P^θ is also supported on $\mathcal{W}$, then

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

where f^θ denotes the density of P^θ; 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.[7] 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 $t \to \infty$.[8] The KLDR is also tightly linked to asymptotics of Bayesian learning, as discussed in the following subsection.

Model $\theta \in \Theta_d$ is a *pseudo-true* d-state model if $\text{KLDR}(\theta) \le \text{KLDR}(\tilde\theta)$ for all $\tilde\theta \in \Theta_d$. If the agent's set of models contains a model θ such that $f^\theta(y_1,\ldots,y_t) = \mathbb{f}(y_1,\ldots,y_t)$ almost everywhere and for all t, then any pseudo-true d-state model is observationally equivalent to the true process.[9] The set of models Θ_d 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 of models is misspecified. The following proposition states that the pseudo-true models are observationally equivalent to the true process when the set of models is correctly specified:

Proposition 1. *Suppose the set Θ_d of d-state models is correctly specified. Then any pseudo-true d-state model P^θ 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 statement is 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 corresponds to using a parsimonious time-series model to capture the essential features of

[6] The 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.

[7] See, for instance, Hansen and Sargent (2008).

[8] See, for instance, Shalizi (2009).

[9] Processes P and $\tilde P$ are *observationally equivalent* if all their finite-dimensional marginal distributions are identical.

a large data set. Unless otherwise specified, I assume throughout the paper that $d \leq n$. However, 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 evolve over time. 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 $t \to \infty$, the agent's posterior will assign a probability of one to the set of pseudo-true models.[10]

This result offers an "as if" interpretation of the pseudo-true d-state models. One can assume that the 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, as long as the agent's prior is supported on the set of d-state models, she 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.[11]

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.

[10] While 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.

[11] One 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 the pseudo-true d-state models. See, for instance, Theorem 2 of Douc and Moulines (2012).

Agents' use of pseudo-true models should therefore be viewed 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 Models

In this section, I characterize the set of pseudo-true d-state models starting with the $d = 1$ case. As a preliminary step, I discuss a useful property of the pseudo-true models, which is of independent interest.

3.1 The Invariance Property

I begin with a result that shows the invariance of the pseudo-true d-state models to linear transformations of the observable. Consider an agent who, instead of observing vector $y_t \in \mathbb{R}^n$, observes vector $\tilde{y}_t = Ty_t \in \mathbb{R}^m$, where T denotes an $m \times n$ matrix. As long as T is a rank-n matrix, y_t and $\tilde{y}_t$ convey the exact same information. Thus, one might expect that the agent's beliefs when she observes y_t are consistent with her beliefs when she instead observes $\tilde{y}_t$.

The following definition formalizes the notion that two probability distributions are consistent with each other given a linear transformation of the observable. Let $T \in \mathbb{R}^{m\times n}$ be a matrix and P be a probability distribution over infinite sequences in $\mathbb{R}^n$. The probability distribution over infinite sequences in $\mathbb{R}^m$ induced by T and P is denoted by $T(P)$ and defined as $T(P)(\mathcal{Y}) \equiv P\left(\{y_t\}_{t=-\infty}^{\infty} : \{Ty_t\}_{t=-\infty}^{\infty} \in \mathcal{Y}\right)$ for any measurable set $\mathcal{Y} \subseteq \mathbb{R}^{m\mathbb{Z}}$.[12] If the observable y_t follows the stochastic process $\mathbb{P}$, then its linear transformation $\tilde{y}_t = Ty_t$ follows the transformed process $T(\mathbb{P})$. The following result establishes that transforming the observable by a rank-n matrix leads the set of pseudo-true models to be transformed accordingly:

Theorem 1 (linear invariance)**.** *Suppose $T \in \mathbb{R}^{m\times n}$ is a rank-n matrix. Then P^θ is a pseudo-true d-state model given true model $\mathbb{P}$ if and only if $T\left(P^\theta\right)$ is a pseudo-true d-state model given true model $T(\mathbb{P})$.*

The result shows that an agent using simple models can discern all linear intratemporal relationships among the observables while facing significant constraints in understanding complex intertemporal relationships. While arguably stark, this dichotomy highlights the paper's premise that forecasting is challenging because it requires forecasters to recognize stochastic patterns that unfold over time. The result makes it possible to abstract from the cognitive costs of acquiring information about a large cross-section of variables and the mistakes individuals make when dealing with cross-sectional complexity, allowing me to instead concentrate on time-series complexity.

[12] The probability distribution induced by a mapping is formally known as the *pushforward measure*.

The linear invariance property makes the predictions of the framework invariant to the exact specification of the variables included in the vector of observables. The agent's pseudo-true models and forecasts only depend on the observables' information content, not on how that information 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 to how she forms her expectations. Likewise, the agent's expectations remain unchanged if the vector of observables is augmented with linear combinations of variables already in her information set.

The theorem thus suggests that it is without loss to assume that the vector of observables is free of redundant variables. Define the lag-l autocovariance matrix of the true process as follows:

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

When y_t includes redundant variables, the variance-covariance matrix Γ_0 is singular, and the true process is degenerate.[13] In such cases, a lower-dimensional vector $\tilde{y}_t$ and a full-rank matrix T exist such that $\mathbb{E}[\tilde{y}_t\tilde{y}_t']$ is non-singular and $y_t = T\tilde{y}_t$. Therefore, by Theorem 1, the pseudo-true models given y_t can be found by first finding the pseudo-true models given $\tilde{y}_t$ and then applying transformation T. This observation implies that there is no loss of generality in assuming that the variance-covariance matrix Γ_0 is non-singular and that the agent only considers subjective models with non-singular variance-covariance matrices.[14] I maintain these assumptions throughout the rest of the paper.

3.2 Pseudo-True One-State Models

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

The agent's pseudo-true one-state forecasts turn out to depend on the true process only through the unconditional variance and the autocorrelation structure of the vector of observables. The autocorrelations are measured by a novel set of objects, which I refer to as autocorrelation matrices. I define the lag-l *autocorrelation matrix* of the observable under the true process as follows:[15]

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

The concept of autocorrelation matrices naturally extends the idea of autocorrelation functions. If the observable y_t is a scalar, C_l simplifies to the standard autocorrelation function at lag l.

[13] A probability distribution on a space is said to be *degenerate* if it is supported on a manifold of lower dimension.

[14] 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 the 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.

[15] Here and throughout the paper, I follow the usual convention that, for a symmetric positive definite matrix X, the square-root matrix $X^{\frac{1}{2}}$ is the unique symmetric positive definite matrix that satisfies $X^{\frac{1}{2}}X^{\frac{1}{2}} = X$.

However, when the observable is an n-dimensional vector, 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. When the spectral radius of C_l is close to zero for all l, the process is close to being i.i.d., whereas when the spectral radius of C_l is close to one, then the process is close to being unit root.[17]

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

Theorem 2. *Under any pseudo-true one-state model θ, the agent's s-period-ahead forecast is given by*

$$E_t^{\theta}[y_{t+s}] = a^s(1-\eta)qp' \sum_{\tau=0}^{\infty} a^{\tau}\eta^{\tau} y_{t-\tau}, \tag{4}$$

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 $u'u = 1$.

The endogenous variables a, η, p, and q have intuitive meanings. The scalar a represents the persistence of the subjective latent state. If $a = 0$, the subjective state is i.i.d., whereas if $a = 1$, it follows a unit-root process.[18] The scalar η captures the perceived 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.

It follows standard Kalman filter results that the agent's forecasts take the form of equation (4) for *some* a, η, p, and q. The substance of the result is rather characterizing the (a, η, p, q) tuple that lead to a model with minimal KLDR from the true process. The theorem suggests a tractable

[16]See Lemma D.2 in the Online Appendix for a proof.
[17]The spectral radius $\rho(X)$ of matrix X denotes the maximum among the magnitudes of eigenvalues of X.
[18]The theorem does not rule out the possibility that $|a| = 1$, in which case the corresponding state-space model might not be stationary ergodic. However, Lemma D.3 in the Online Appendix establishes that any pseudo-true one-state model inherits the stationarity and ergodicity of the true process.

way of computing the pseudo-true one-state forecasts in any stationary and ergodic environment given only the knowledge of the true autocorrelation matrices.

The theorem significantly reduces the computational complexity of finding the set of pseudo-true models. It concentrates out all parameters in the agent's models except for two scalars. As a result, the optimization problem simplifies from a problem over a $2n$-dimensional non-compact manifold to a much simpler problem over a two-dimensional compact rectangle.[19] Furthermore, 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 next result characterizes the perceived variance-covariance matrix of the observable under the pseudo-true one-state models:

Theorem 3. *Given any pseudo-true one-state model θ, the subjective variance-covariance of the vector of observables, $E^{\theta}[y_t y_t']$, coincides with the true variance-covariance matrix, $\Gamma_0 \equiv \mathbb{E}[y_t y_t']$.*

The theorem hinges on two main assumptions: First, there are no constraints on the agent's set of models other than the bound on the number of subjective state variables. Put differently, matrices A, B, Q, and R of representation (1) are unrestricted other than the constraint on their dimension. This flexibility allows the agent to represent any cross-sectional correlation pattern by an appropriate selection of matrices A, B, Q, and R. Second, the agent uses a model that minimizes the KLDR from the true process. This leads her to a set of such matrices that perfectly capture the true cross-sectional correlations.

Theorems 2 and 3 fully characterize the pseudo-true one-state models in terms of the true variance-covariance matrix Γ_0 and the tuple (a, η, p, q), which in turn only depends on the true autocorrelation matrices $\{C_l\}_{l=1}^{\infty}$. Any unconditional or conditional moment of the pseudo-true one-state model can, in turn, be found in terms of Γ_0 and (a, η, p, q).

3.3 Pseudo-True One-State Models Under Exponential Ergodicity

The pseudo-true one-state models can be found in closed form given a class of true stochastic processes that naturally arise in applications. 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 $l \geq 1$, where $\rho(C_l)$ denotes the spectral radius of C_l.

Exponential ergodicity is stronger than ergodicity. Ergodicity requires that the serial correlation at lag l decays to zero as $l \to \infty$. Exponential ergodicity requires the rate of decay to be faster than $\rho(C_1)$. Although exponentially-ergodic processes only constitute a subset of the class of stationary ergodic processes, many standard processes are exponentially ergodic. For instance, the vector of observables follows an exponentially-ergodic process if it is a spanning linear combination of n independent AR(1) shocks.

[19] The set of all d-state models is a non-compact manifold of dimension $2nd$ (Gevers and Wertz, 1984). Additionally, the KLDR is a non-convex function of $\theta = (A, B, Q, R)$.

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 4. *Suppose the true process is exponentially ergodic. Under any pseudo-true one-state model θ, the agent's s-period-ahead forecast is given by*

$$E_t^\theta[y_{t+s}] = a^s q p' y_t, \tag{5}$$

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 $q = \Gamma_0^{\frac{1}{2}} u$.

A remarkable feature of the characterization in Theorem 4 is that the agent's forecasts only depend on the last realization of the observable (and not its lags). In other words, the pseudo-true one-state model is *Markovian* if the true process is exponentially ergodic. This property might come as a surprise in light of the fact that in the correctly-specified case forecasts obtained using the stationary Kalman filter generically use the entire history of the observable. The seeming discrepancy between the two results is due to misspecification of the agent's set of models in Theorem 4, as illustrated by the following example:

Example 1. Suppose the observable is scalar and follows an AR(∞) process: $y_{t+1} = \sum_{\tau=1}^{\infty} \phi_\tau y_{t+1-\tau}$.[20] It is then immediate that the one-step-ahead forecast of the observable under the true, correctly-specified model is given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \phi_\tau y_{t+1-\tau}.$$

Contrast this with what an agent can do when she is constrained to use (misspecified) one-state models. Under any such model θ, the agent's one-step-ahead forecast takes a similar form:

$$E_t^\theta[y_{t+1}] = \sum_{\tau=1}^{\infty} \alpha_\tau y_{t+1-\tau}.$$

However, the restriction to one-state models constrains coefficients $\{\alpha_\tau\}_{\tau=1}^{\infty}$ to be given by $\alpha_\tau = (1-\eta)a^\tau \eta^{\tau-1}$ for some $a \in [-1, 1]$, some $\eta \in [0, 1]$, and all τ. Therefore, the pseudo-true one-state model is the model that picks $\{\alpha_\tau\}_{\tau=1}^{\infty}$ to minimize the KLDR subject to the constraint that $\alpha_\tau = (1-\eta)a^\tau \eta^{\tau-1}$ for all τ. The agent wants to set α_τ to a value that is related to the correlation of y_{t+1} and $y_{t-\tau}$, but the constraint prevents her from fine-tuning the $\{\alpha_\tau\}_{\tau=1}^{\infty}$ coefficients. When the true process is exponentially ergodic, y_{t+1} is much more correlated with y_t than it is with lags of y_t. Then, the best such a constrained agent can do is to fine-tune the coefficient of y_t and entirely disregard its lags. In other words, the constrained minimizer of the KLDR is Markovian even though the unconstrained minimizer is not.

The next example illustrates the use of Theorem 4 in the context of a commonly-used process:

[20] Such a representation exists for generic processes in the class of mean-zero, purely non-deterministic, and stationary processes.

Example 2. Suppose the true process $\mathbb{P}$ has the following representation:

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

where $\epsilon_t \sim \mathcal{N}(0, \Sigma)$,

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

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

$H \in \mathbb{R}^{n\times n}$ is an invertible square matrix, and $1 > |\alpha_1| > |\alpha_2| > \cdots > |\alpha_n| > 0$. 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 4 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 $q = (H'VH)^{\frac{-1}{2}} H' V e_1$, where $V \equiv (I - F^2)^{-1}\Sigma$ is the variance-covariance matrix of f_t and e_1 denotes the first coordinate vector.[21]

The agent's forecasts take a particularly simple form when 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

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

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.

A noteworthy feature of the pseudo-true model in Example 2 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 small 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 the subjective latent state in the agent's pseudo-true model. Therefore, the persistence parameter cannot depend on the volatilities.

The example also 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

[21] See the proof of Lemma D.5 for a derivation.

expectations but misses the dynamics of the other components. The intuition for the result is easiest to see when the most persistent component is close to being unit root. In that case, poorly tracking the most persistent component 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 KLDR minimization. Therefore, any pseudo-true model tracks the component close to unit root as best possible, even if doing so results in errors in forecasting the other components. In Section 4.1, I generalize the insight of this example by formally establishing persistence bias in a more general context.[22]

Example 2 can be generalized by relaxing the assumption that matrices F and Σ are diagonal and allowing for non-Gaussian innovations. The key requirement for the process to be exponentially ergodic is that matrix H in representation (6) is full rank. This assumption can be seen as a full-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. Appendix B discusses this generalization in further detail.

3.4 Pseudo-True d-State Models

I end this section by discussing how the insights from the $d = 1$ case generalize when $d > 1$. To characterize the pseudo-true d-state models, one needs to find models $\theta = (A, B, Q, R)$ that minimize the KLDR from the true process. Doing so requires minimizing a non-convex function over a non-compact set, consisting of all the matrices A, B, Q, and R of appropriate dimensions. This problem does not lend itself to an analytical solution without further restrictions.

I proceed by restricting the models the agent considers to be Markovian. A d-state model θ is *Markovian* if P^θ satisfies the Markov property, i.e., $P^\theta(y_{t+1}|y_t, y_{t-1}, \dots) = P^\theta(y_{t+1}|y_t)$. An agent who believes the observable follows a Markovian d-state model believes (a) that the current realization of the observable contains all the information required for forecasting and (b) that all the relevant information contained in y_t can be summarized by a d-dimensional state variable. The following proposition provides a necessary and sufficient condition for a model to be Markovian:

Proposition 2. *Let $Var_t^\theta(y_{t+1})$ denote the variance-covariance matrix of y_{t+1} given model θ and conditional on the history $\{y_\tau\}_{\tau \le t}$ of the observable, and let $Var^\theta(y_{t+1}|z_t)$ denote the corresponding variance-covariance matrix conditional on the time-t realization of the subjective latent state. $Var_t^\theta(y_{t+1}) \ge Var^\theta(y_{t+1}|z_t) = B'QB + R$ for any d-state model θ, with equality if and only if θ is Markovian.*[23]

[22] Bidder 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.

[23] I use the usual convention that $X \ge Y$ for symmetric positive semidefinite matrices X and Y if $X - Y$ is positive semidefinite.

The proposition highlights an intuitive property of Markovian models. Note that the agent can observe the history $\{y_\tau\}_{\tau\leq t}$ of the observable but not the subjective latent state z_t. The first part of the proposition shows that the agent cannot forecast any better than if she knew the realization of the latent subjective state. In other words, the forecast error given z_t provides a lower bound on the forecast error given $\{y_\tau\}_{\tau\leq t}$. The second part of the proposition shows that the agent can achieve this lower bound when her model of the world is Markovian. She can then forecast as well as an agent who knows the latent subjective state, because all the relevant information in the latent state can be extracted from the realized history of the observable. Markovian models can thus be seen as models that feature *full information*.

Markovian models constitute only a subset of the class of all state-space models of a given dimension. However, the pseudo-true one-state models happen to be Markovian when the true process is exponentially ergodic, as shown by the following corollary of Theorem 4:

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

The result shows that constraining the agent to Markovian models is without loss when $d = 1$ and the true process is exponentially ergodic. Even with the flexibility to choose non-Markovian models, an agent who is attempting to minimize the KLDR from an exponentially ergodic process settles on a Markovian model. Whether this result continues to hold for d-state models with $d > 1$ remains an open question. However, I can still make progress by taking the restriction to Markovian models as an assumption and characterizing the resulting pseudo-true models. A Markovian d-state model θ is a *pseudo-true Markovian d-state model* if $\mathrm{KLDR}(\theta) \leq \mathrm{KLDR}(\tilde{\theta})$ for any Markovian d-state model $\tilde{\theta}$.

The pseudo-true Markovian models have a number of appealing properties. They satisfy a version of the linear-invariance result of Theorem 1. They have similar Bayesian and quasi-maximum-likelihood learning foundations as other pseudo-true models. Perhaps most importantly, they can be fully characterized in closed-form in some useful cases:

Theorem 5. *Suppose either $d = 1$ or the lag-one autocovariance matrix is symmetric. Then the following statements hold:*

(a) Under any pseudo-true Markovian d-state model θ, the agent's s-period-ahead forecast is given by

$$E_t^\theta[y_{t+s}] = \sum_{i=1}^{d} {a_i}^s q_i {p_i}' y_t, \tag{7}$$

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

(b) Under any pseudo-true Markovian d-state model θ, the subjective variance-covariance of the vector of observables, $E^{\theta}[y_t y_t']$, coincides with the true variance-covariance matrix, $\Gamma_0 \equiv \mathbb{E}[y_t y_t']$.

The result shows that the insights from the analysis of one-state simple models broadly carry over to d-state ones. In particular, agents who are restricted to Markovian d-state models exhibit a form of persistence bias. They focus on perfectly forecasting the d most persistent components of the vector of observables at the expense of the other components. Moreover, agents who are constrained to use Markovian d-state models uncover the true variance-covariance matrix of the observable.

The theorem also 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.[24] However, Theorem 5 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. Lastly, Theorem 5 suggests that the PCA estimator is consistent (at least under the theorem's assumptions) even if the number of states is misspecified.[25] 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.

4 Behavioral Implications

In this section, I apply the characterization results from the previous section to develop the behavioral implications of the simple models framework. Throughout the section, I maintain the assumption that at least one of the following is satisfied for every agent: (a) the agent is constrained to use one-state models and the true process is exponentially ergodic; (b) the agent is constrained to use Markovian one-state models; or (c) the agent is constrained to use Markovian d-state models and the lag-one autocovariance matrix, Γ_1, is symmetric.

To flesh out the behavioral implications of the framework, I embed it in a reduced-form economy. Consider a finite set of agents, indexed by $j = 1, \ldots, J$. In every period t, each agent j takes a purely forward-looking decision x_{jt}, which depends on her forecasts via the best-response function

$$x_{jt} = E_{jt}\left[\sum_{s=1}^{\infty} c_{js}' y_{t+s}\right], \tag{8}$$

[24] See, for instance, Stock and Watson (2002).

[25] An estimator for a misspecified model is consistent if the estimate converges to a pseudo-true model as the sample size goes to infinity.

where $y_t \in \mathbb{R}^n$ is as before the vector of observables, $E_{jt}[\cdot]$ denotes agent j's subjective forecasts, and $c_{js} \in \mathbb{R}^n$ are preference parameters satisfying $\sum_{s=1}^{\infty} \|c_{js}\|_2 < \infty$ for all j.[26] I continue to take the true process $\mathbb{P}$ as a primitive of the economy and assume that agent j can only entertain state-space models with no more than d_j states. In Appendix C, I provide an analysis suggesting that the partial equilibrium insights would generalize to a general equilibrium economy, in which $\mathbb{P}$ itself is an endogenous outcome of agents' choices.

The reduced-form specification in (8) allows the derivation of sharp theoretical results, which highlight the role of simple models and biased forecasts and are independent of the specifics of agents' decision problems. These results are valid up to first order for purely forward-looking decisions that depend non-linearly on the forecasts of the observable. They also hold 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 general equilibrium macro models.

4.1 Persistence Bias

Decomposing the observable into its more and less persistent components will be useful for the subsequent discussions:

Proposition 3. *Let a_i denote the ith largest eigenvalue of the first autocorrelation matrix, C_1, in magnitude, and let u_i denote the corresponding eigenvector, normalized such that $u_i' u_k = \mathbb{1}_{\{i=k\}}$ for all i and k. The observable can be decomposed as follows:*

$$y_t = \sum_{i=1}^{n} y_t^{(i)} q_i, \tag{9}$$

where $y_t^{(i)} \equiv p_i' y_t$, $p_i \equiv \Gamma_0^{\frac{-1}{2}} u_i$, $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$, u_i is as in Theorem 5, and scalars $y_t^{(i)}$ all have unit variance. If ρ_i denotes the lag-one autocorrelation of $y_t^{(i)}$, then $|\rho_1| \geq |\rho_2| \geq \cdots \geq |\rho_n|$.

This proposition represents the observable in terms of the basis vectors $\{q_i\}_{i=1}^n$, with $y_t^{(i)}$ denoting the components (or coordinates) of y_t with respect to this basis. The components of y_t are sorted by their persistence, with $y_t^{(1)}$ representing the *most persistent component* and $y_t^{(n)}$ the *least persistent component* of the observable. This decomposition is valid for arbitrary stationary stochastic processes and is independent of agents' forecasting and decision problems. However, the way agents' choices respond to changes in the observable neatly aligns with the decomposition in (9). This is shown in the following corollary of Theorems 4 and 5:

Corollary 2 (persistence bias)**.** *Agent j's time-t forecasts and forward-looking actions only respond to changes in the d_j most persistent components of y_t.*

[26] The assumption that each agent takes a single action is without loss of generality. The analysis would be identical if one instead assumed that agent j makes multiple choices in each period, with the kth action of agent of j given by $x_{jkt} = E_{jt}\left[\sum_{s=1}^{\infty} c_{jks}' y_{t+s}\right]$.

Agents who use pseudo-true d-state models treat the more and less persistent components of y_t in qualitatively different ways. A change in the current value of the observable can be decomposed into changes in the components $y_t^{(i)}$ of y_t. Agents do not change their forecasts in response to changes in the least persistent components of y_t. Consequently, their forward-looking actions also remain unresponsive to changes in these less persistent components.

It is worth noting that agents' forecasts and actions are unresponsive to changes in the less persistent components of the observable only on impact. In general, different components of y_t do not evolve independently. Therefore, a change in the current value of $y_t^{(i)}$ could lead to changes in the values of $y_{t+s}^{(j)}$ for some $j \neq i$ and $s > 0$. This can result in a delayed response of agents' forecasts and actions to changes in the observable's less persistent components.

4.2 Increased Comovement

Constraining agents to use simple models increases the comovement between their forward-looking choices. The argument for this prediction is best seen by considering agents j and k, both of whom are constrained to use one-state models. Because of persistence bias, the agents' time-t actions can be written as time-invariant linear functions of the observable's most persistent component. More specifically, $x_{jt} = g_j^{(1)} y_t^{(1)}$ and $x_{kt} = g_k^{(1)} y_t^{(1)}$, where $g_j^{(1)}$ and $g_k^{(1)}$ are constants that depend on the true process and the agents' preferences. Thus, one agent's actions can be expressed as a constant multiple of the other agent's actions. In other words, the agents' actions comove perfectly. The following proposition formalizes and extends this conclusion:

Proposition 4. *Let $D \equiv \max_j d_j$ denote the largest value of d_j among agents and $x_t \equiv (x_{jt})_j \in \mathbb{R}^J$ denote the vector containing agents' time-t actions. Given generic true processes, x_t has the factor structure*

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

where G is a $J \times D$ matrix of loadings and $y_t^{(1:D)}$ is the D-dimensional vector consisting of the D most persistent components of the observable.[27]

The proposition establishes that the J-dimensional vector of all the forward-looking actions of all the agents in the economy moves with the D factors collected in $y_t^{(1:D)}$. The number of factors depends solely on the complexity of agents' models, while the composition of these factors depends on the properties of the true process. The loadings of actions on different factors depend on the preference parameters c_{js}. If D is much smaller than J (as is often reasonable to assume to be the case), then a large number of actions comove with movements in a small number of factors.

Agents' actions exhibit comovement, both across agents with the same value of d and across those who use models of varying dimensions. To see the intuition for this result, consider agents

[27]The result requires all pseudo-true models of a given dimension to be observationally equivalent, a condition that holds for generic true processes.

j and k who use models of dimensions d_j and $d_k > d_j$, respectively. While the agents disagree on the number of state variables needed to forecast the observable, they agree on what d_j of those state variables ought to be. The d_j states used by agent j are a subset of the d_k states used by agent k (up to linear transformations). This strong form of comovement is a unique prediction of the framework of simple models. It relies on the fact that pseudo-true d-state models rank the components of y_t consistently across d: As d increases, pseudo-true forecasts condition on additional components of y_t but without altering the components already being used to forecast.

A low-dimensional factor structure is one natural expression of comovement. Another commonly used comovement measure is the Pearson correlation coefficient between two variables. The following corollary of Proposition 4 shows that constraining any two agents to one-state models increases the correlation between their actions: [28]

Corollary 3. *Consider actions j and k, both of the form (8), taken by agents j and k with $d_j = d_k = 1$. Generically,*

$$1 = \left|Corr\left(x_{jt}^{1d}, x_{kt}^{1d}\right)\right| > \left|Corr\left(x_{jt}^{RE}, x_{kt}^{RE}\right)\right|,$$

where x_{jt}^{1d} and x_{jt}^{RE} denote agent j's time-t action when using a pseudo-true one-state model and the true model, respectively.

The time-t actions of agents using a pseudo-true one-state model depend solely 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 holds regardless of the specifics of preferences, technology, or market structure. It holds both in partial equilibrium and in general equilibrium, as suggested by the analysis in Appendix C. Angeletos, Collard, and Dellas (2020) find that a "main business cycle shock" explains the bulk of movements in macroeconomic aggregates at business cycle frequencies. The above analysis suggests that there is always a main shock—as long as decisions are sufficiently forward looking and agents use simple models. It also shows that the main shock is an endogenous index whose composition depends on the primitives of the economy, the stochastic properties of the shocks that hit it, and the parameters of policy rules.

4.3 Under- and Over-Extrapolation

The framework proposed in this paper is neither a model of under-extrapolation nor of over-extrapolation. Instead, agents who use simple models forecast using a parsimonious model that provides an approximation to the true process and balances forecast errors across different horizons and variables. Consequently, simple models do not lead to mistakes that invariably go in the same direction. In fact, agents who use a pseudo-true d-state model under-extrapolate some variables and over-extrapolate others.

[28] This corollary does not generalize beyond the one-state case. Constraining agents to models with $d_j, d_k > 1$ states might *decrease* the correlation between their actions relative to the rational-expectations benchmark.

Proposition 5. *Let $y_t^{(1)}$ denote the most persistent component of y_t and $y_t^{(n)}$ denote its least persistent component. If the true process is exponentially ergodic and $d_j < n$, then:*

(a) Agent j overestimates the magnitude of $y_t^{(1)}$'s autocorrelation at all lags.

(b) Agent j underestimates the magnitude of $y_t^{(n)}$'s autocorrelation at all lags.

The following example illustrates the result:

Example 3. 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

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

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 ith 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 ith 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 component in the standard coordinates and its least persistent component is its second component in the standard coordinates. The spectral radius of the lag-l autocorrelation matrix satisfies

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

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 4. 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 who uses 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.

Agents who use simple models over-extrapolate from changes in the most persistent components of the observable and under-extrapolate from changes in the least persistent ones. For observables with intermediate persistence, the pattern could be under- or over-extrapolation depending on the variable and the horizon being considered. These predictions set this paper's framework apart from those that hardwire under- or over-extrapolation.

5 Application to the New-Keynesian Model

As the first application of the general framework, I study the standard three-equation new-Keynesian model.[29]

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).[30]

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

[29]Technically 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 nominal interest rate is set (e.g., a Taylor rule), I model the interest rate as a random variable with arbitrary dependence on the current realizations and lags of the output gap and inflation rate.

[30]The Taylor principle is not necessary for equilibrium uniqueness in my setup. The sunspot equilibria in the new-Keynesian model require agents to coordinate on payoff-irrelevant 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.

market clearing conditions but not rational expectations.[31] The second step is to supplement the temporary equilibrium with the model of expectation 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.[32] These temporary-equilibrium conditions are given by

$$\hat{x}_t = -\sigma\left(\hat{i}_t - r_t^n\right) + E_{ht}\left[\sum_{s=1}^{\infty}\beta^s\left(\frac{1-\beta}{\beta}\hat{x}_{t+s} - \sigma\left(\hat{i}_{t+s} - r_{t+s}^n\right) - \frac{\sigma}{\beta}\hat{\pi}_{t+s}\right)\right], \tag{10}$$

$$\hat{\pi}_t = \kappa\hat{x}_t + \mu_t + E_{ft}\left[\sum_{s=1}^{\infty}(\beta\delta)^s\left(\kappa\hat{x}_{t+s} + \frac{1-\delta}{\delta}\hat{\pi}_{t+s} + \mu_{t+s}\right)\right], \tag{11}$$

where $\hat{x}_t$, $\hat{i}_t$, and $\hat{\pi}_t$ denote the log-deviations of the 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 slope 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_{ht} and E_{ft} denote the subjective expectations of households and firms, respectively.

I assume that the vector $(\hat{i}_t, r_t^n, \mu_t)'$ of the nominal interest rate, technology shock, and cost-push shock follows a mean-zero, stationary, and exponentially ergodic process.[33] This assumption allows me to use Theorem 4 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. Agents' time-t information set is given by the history $\{y_\tau\}_{\tau\le t}$ of vector $y_\tau \equiv (\hat{x}_\tau, \hat{\pi}_\tau, \hat{i}_\tau, r_\tau^n, \mu_\tau)'$, consisting of the 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 that agents are constrained to use one-dimensional state-space models of the form (1) to forecast y.

[31] The 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.

[32] The details of this derivation can be found, among other places, in Angeletos and Lian (2018) and Gáti (2020).

[33] The new-Keynesian literature often assumes that the nominal interest rate follows a Taylor rule, which sets the rate as a linear function of the 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 estimated 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.

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 the *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.[34] 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_t that spans the subspace spanned by y_t and then using the linear-invariance result to find their forecasts of y_t.

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 simplifies the algebra involved in finding the equilibrium considerably. Second, it makes the estimation of the model more straightforward. By Theorem 4, 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 the output gap, inflation rate, and nominal interest rate, those autocovariance matrices have readily available empirical counterparts.

The following proposition summarizes the equilibrium characterization:

Proposition 6. *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 $\tau \le t$. In the linear constrained-rational-expectations equilibrium,*

$$\hat{x}_t = \frac{1}{1 - p_x\gamma_x - p_\pi(\gamma_\pi + \kappa\gamma_x)}\left[\gamma_x(p_i\hat{i}_t + p_\pi\mu_t) - \sigma(1 - \gamma_\pi p_\pi)(\hat{i}_t - r_t^n)\right], \tag{12}$$

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

where

$$\gamma_x \equiv a(q_x - \sigma q_\pi), \tag{14}$$

$$\gamma_\pi \equiv a\beta q_\pi, \tag{15}$$

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 $q \equiv (q_x, q_\pi, q_i)' \equiv \Gamma_0^{\frac{1}{2}} u$.

[34] The 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.

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 the 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 the autocovariance matrices of f_t and the 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. 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 $\zeta \in \{x, \pi, i\}$, 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 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, agents' expectations of ζ are well anchored; they do not respond to innovations in any of the observables.

The framework can be used to study optimal 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 7 (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 $y_\tau \equiv (\hat{x}_\tau, \hat{\pi}_\tau, \hat{i}_\tau, r^n_\tau, \mu_\tau)'$ 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., $\hat{i}_t = r^n_t$, 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 agents' pseudo-true one-state expectations closer to rational expectations, thus making the constraint on the complexity of agents' models non-binding. As a result, the economy inherits the divine coincidence property from the rational-expectations version of the model. In equilibrium, $q_x = q_\pi = 0$; that is, agents' expectations of output and inflation are perfectly anchored.

Limits to the complexity of agents' models do 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 effects 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 their 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 $\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$ going forward. The announcement augments agents' time-t information set to include $\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$ (in addition to $\{y_\tau\}_{\tau \le t}$). Therefore, agents' time-t forecasts under forward guidance are the conditional expectations $E^{\theta^*}_{t,\mathrm{FG}(T)}[\cdot] \equiv E^{\theta^*}_t[\cdot|\{y_\tau\}_{\tau\le t}, \hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}]$, where θ^* denotes the pseudo-true one-state model agents use in the equilibrium of the economy. But agents' forecasts are Markovian by Theorem 4 and the assumption that the true process is exponentially ergodic. Therefore, $E^{\theta^*}_{t,\mathrm{FG}(T)}[\cdot] = E^{\theta^*}_t[\cdot|y_t, \hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}]$.

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 $\zeta \in \{\hat{x}, \hat{\pi}, \hat{\imath}, r^n, \mu\}$

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

where $\omega_T \equiv (\zeta_t, \hat{i}_{t+1}, \ldots, \hat{i}_{t+T})'$, $\Sigma_{\zeta_s\omega_T} \equiv E^{\theta^*}[\zeta_{t+s}\omega_T']$, and $\Sigma_{\omega_T\omega_T} \equiv E^{\theta^*}[\omega_T\omega_T']$. Note that the covariance matrices that show up in agents' forecasts of ζ are subjective covariance matrices which depend on agents' subjective model. But the subjective model is just the pseudo-true one-state model, which is fully characterized by Proposition 6.

The response of the economy to forward guidance can be computed analytically. Substituting for agents' forecasts in (10) and (11) and simplifying the resulting expression, I obtain

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

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

The ν's have intuitive interpretations: ν_{xi} and $\nu_{\pi i}$ are the current interest-rate elasticities of output and inflation, respectively, whereas ν_{xi_s} and $\nu_{\pi i_s}$ are the elasticities of output and inflation with respect to the s-period-ahead guidance about the path of the 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 ν's are rather cumbersome, so I calibrate the model and numerically study the effects of forward guidance.

5.5 Quantification

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 the 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 the 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 E.2 for the full estimated autocovariance matrices at lags zero and one, where I also verify that the estimated process is indeed exponentially ergodic.

There are no free parameters for agents' expectations (other than d, which I have set equal to one). 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 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 as highly persistent but not unit root. Their estimate of the state is much more sensitive to changes in inflation than to output or interest rate. High output makes agents optimistic about the subjective state, while high inflation and high interest rate make them pessimistic. Finally, agents' forecasts of inflation and interest rate move much more with changes in their estimate of the state than do their forecasts of 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-step-ahead forecast of the 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 Quantitative 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 responses at the time of announcement as the duration of guidance T is varied. The response of output to the rate cut accompanied by a promise to keep the rate low for another quarter is 48% higher than the response to a rate cut without any guidance; the corresponding figure for inflation is 17%. 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% and *decreases* the response of inflation 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.

Agents' misperception of the intertemporal relationships among aggregate variables moderates the power of forward guidance. Agents use a simplified statistical model that is fit to their past observations to make an out-of-sample forecast in a new regime with forward guidance. This

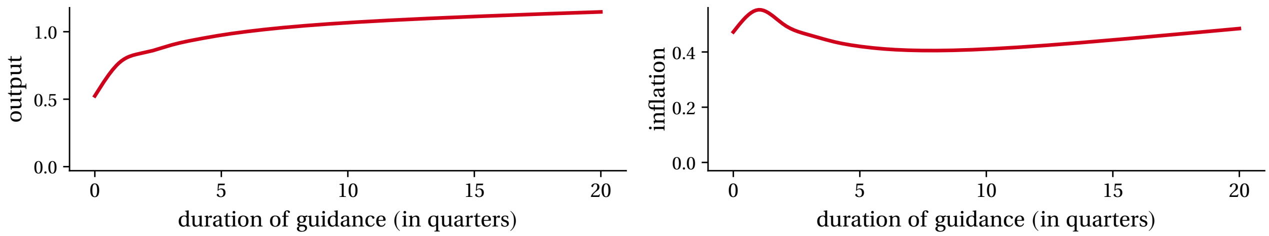

Figure 1. The power of forward guidance

Notes: This figure plots the responses of the output gap and inflation rate to a 100 basis point cut in the current nominal interest rate combined with an announcement by the central bank that the nominal rate will be kept at -1% for an additional T quarters. The responses are plotted at the time of announcement as the duration of guidance T is varied. The output gap is measured in percentage differences between Real GDP and Potential Output. The inflation rate is measured in percentage changes in GDP Deflator.

estimated pseudo-true model understates the extent of serial correlation in the observables relative to the rational-expectations benchmark. For instance, under the 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. 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 $\hat{x}_{t-s}$ to be 0.74×0.985^s. Therefore, changes in the interest rate in period $t+T$ have a modest impact on agents' time-t forecasts of the output gap and inflation rate in periods $t+1, t+2, \ldots, t+T-1$, and the impact decays exponentially with T.

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

$$\hat{a}_t = \rho \hat{a}_{t-1} + \epsilon_t. \tag{16}$$

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 interest rate, and w_t is the wage rate. Market clearing determines investment:

$$i_t = o_t - c_t.$$

6.2 Equilibrium

I find the equilibrium under the assumption that households know the steady-state of the economy and use a one-state model to represent the fluctuations of aggregate variables around the steady state. Specifically, I assume that households 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. However, instead of having rational expectations, households believe that y_t follows a model of the form (1) with $d = 1$ and use a pseudo-true one-state model to forecast future values of y.

The equilibrium definition is as in the new-Keynesian application. A constrained-rational-expectations 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 decisions given their subjective model θ^*; and (ii) θ^* is a pseudo-true one-state model given the stochastic process $\mathbb{P}^*$. In Online Appendix F, I provide a more formal definition of equilibrium and discuss how one can compute it.

6.3 Quantification

The model is 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 normalize the standard deviation of TFP innovations to one.

6.4 Impulse-Response Functions

Figure 2 plots the impulse-response functions (IRFs) to a one-percent increase in TFP. The response of consumption on impact is 83% smaller than the corresponding response under rational expectations (RE). The consumption response continues to be smaller than under RE for six quarters after impact. However, consumption in the economy with boundedly-rational agents eventually overshoots its RE counterpart. The consumption IRF under simple models mimics the IRF in a model with consumption-habit formation, a common feature in the DSGE literature that serves to increase the sluggishness of consumption. Simple models thus provide a novel account of the hump-shaped response of consumption to TFP shocks in empirical studies, which does not rely on auxiliary frictions.[35] [36]

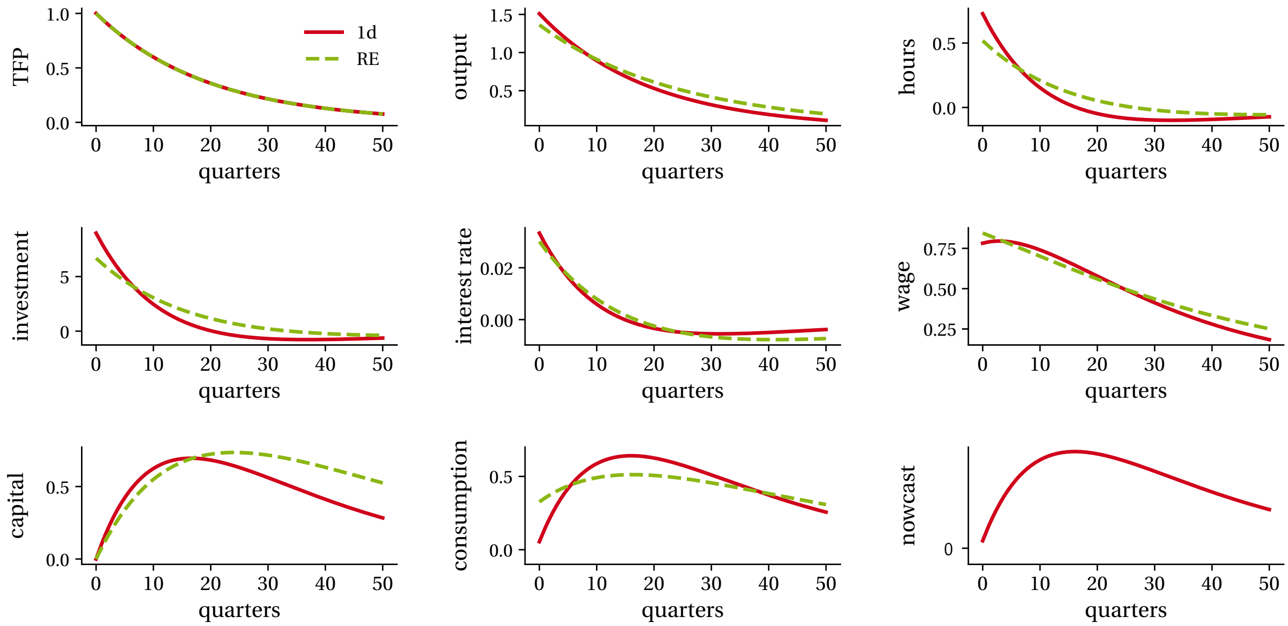

Figure 2. Impulse-response functions to a TFP shock

Notes: This figure plots the impulse-response functions to a one-percent increase in TFP. The variable labeled "nowcast" is defined as households' nowcast $\hat{z}_t$ of the subjective state z_t. Every variable, except for the interest rate and nowcast, is measured in log changes from its steady-state value. The interest rate is measured in percentage-point changes from its steady-state value. The unit of measurement for agents' nowcast is arbitrary.

The key to understanding the consumption IRF is how households form their expectations. Agents' forecasts of aggregate variables are linear functions of their nowcast $\hat{z}_t$ of the subjective state z_t in their pseudo-true one state model. This nowcast is a linear combination of the time-t values of the capital stock and TFP, with the weights determined endogenously in equilibrium:[37]

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

Agents' nowcast is significantly more sensitive to changes in the capital stock than to changes in TFP. This is a manifestation of persistence bias: In equilibrium, the capital stock is more persistent than TFP, a fact that is evident from the impulse-response functions. Therefore, the nowcast moves almost one-for-one with changes in the capital stock. Agents' expectations of every aggregate variable at every horizon is a scaled version of their nowcast. Consequently, they also move in unison with changes in the capital stock.

Consumption inherits the dynamics of agents' expectations. Since the discount factor is close to one in my calibration, consumption is almost purely forward looking. Therefore, it moves in

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

[36] The 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.

[37] Since the scale of z_t is not identifiable either to agents or an outside observer, the scale of $\hat{z}_t$ is intrinsically meaningless. I normalize $\hat{z}_t$ so that $\hat{z}_t = p_k\hat{k}_t + p_a\hat{a}_t$ with $|p_k| + |p_a| = 1$.

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

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

That is, consumption is much more sensitive to changes in agents' nowcast than to current prices and quantities.

The upshot is that consumption comoves, almost perfectly, with the capital stock. The (unconditional) correlation of consumption and the capital stock is 0.999 when households use simple models. This is compared to a correlation of 0.956 under 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 and behaves like a state variable.

The initial underreaction and the subsequent overshooting of consumption increases its unconditional variance by about 10% relative to the RE benchmark. Similarly, hours and investment show increased volatility when agents are constrained to one-state models. They are about 41% and 24% more volatile, respectively, compared to the rational-expectations benchmark. The increased volatility of consumption and hours, resulting from agents' use of simple models, leads to an increase in the cost of business cycles.

I end this section with a note on Angeletos, Huo, and Sastry (2021)'s evidence on the IRFs of subjective expectations. They find that the "consensus forecasts" of various aggregate variables in the Survey of Professional Forecasters exhibit hump-shaped responses to identified aggregate shocks. They also find that the term structure of subjective expectations feature monotone mean-reversion with a persistence parameter that is larger than the true persistence. The behavior of subjective expectations in the RBC model with simple models discussed in this section is consistent with both of these findings.

7 Application to the Diamond–Mortensen–Pissarides Model

For my last application, 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 any realization of labor productivity, 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. An unemployed worker finds a job in period t with probability $p_t = \mu\theta_t^{1-\alpha}$, and a 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. Jobs are 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 of k units of output per period for maintaining a vacancy.

The wage is 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

$$\begin{pmatrix}\hat{a}_t\\ \hat{s}_t\end{pmatrix} = \begin{pmatrix}\rho_a & 0\\ 0 & \rho_s\end{pmatrix}\begin{pmatrix}\hat{a}_{t-1}\\ \hat{s}_{t-1}\end{pmatrix} + \epsilon_t, \tag{17}$$

where $\epsilon_t \sim \mathcal{N}(0,\Sigma)$. This specification allows for labor productivity and the 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 the wage then satisfy the following equations:[38]

$$\theta_t^\alpha = \frac{\mu}{k}(a_t - w_t) + \frac{\mu}{k}E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\prod_{k=1}^{\tau}(1-s_{t+k})(a_{t+\tau}-w_{t+\tau})\right], \tag{18}$$

$$\begin{aligned} w_t = {} & \delta a_t + (1-\delta)b + \delta E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\prod_{k=1}^{\tau}(1-s_{t+k})(a_{t+\tau}-w_{t+\tau})\right] \\ & - (1-\delta)E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\prod_{k=1}^{\tau}(1-s_{t+k}-p_{t+k})(w_{t+\tau}-b)\right]. \end{aligned} \tag{19}$$

The unemployment rate follows the first-order difference equation

$$u_t = u_{t-1} + s_{t-1}(1-u_{t-1}) - \mu\theta_{t-1}^{1-\alpha}u_{t-1}. \tag{20}$$

[38]Nash bargaining only determines the total value delivered to workers and firms and not the timing of 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.

Equations (17)–(20) together with the specification of subjective expectations fully determine the equilibrium. The derivation of these equations and other omitted calculations from this section can be found in Online Appendix G. 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 Equilibrium

I find the equilibrium under the assumption that agents know the steady state of the economy and use a one-state model to represent the fluctuations of aggregate variables around the steady state. I assume that every agent perfectly observes the 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 believe that y_t follows a one-state model of the form (1) and use a pseudo-true 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 agents such that (i) $\mathbb{P}^*$ is derived from 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 G.

7.4 Quantification

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 $\alpha = 0.72$. 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 the separation rate. I normalize the steady-state output per worker to $a = 1$. The flow payoff to workers from unemployment is set to $b = 0.4$.[39]

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 the labor-productivity shock to that of the separation-rate shock to ten. These choices ensure that the (pairwise) correlation coefficients between labor productivity, the 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.

[39] These 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.

7.5 Impulse-Response Functions

The responses of aggregate variables to shocks depend on how agents form their expectations. Agents' expectations of aggregate variables are linear functions of their nowcast $\hat{z}_t$ of the subjective state z_t. The nowcast, in turn, is the following linear combination of the time-t values of the unemployment rate, labor productivity, and the separation rate:[40]

$$\hat{z}_t = -0.848\hat{u}_t + 0.043\hat{a}_t - 0.108\hat{s}_t. \tag{21}$$

The nowcast is almost eight times more sensitive to changes in the unemployment rate than to changes in the separation rate, and it barely responds to changes in labor productivity. This is due to the fact that, in equilibrium, the unemployment rate is more persistent than the two exogenous shocks.

Figure 3 plots the impulse-response functions to a one-percent increase in labor productivity. The responses of the number of vacancies and job-finding rate on impact are 45% smaller than the corresponding responses under rational expectations. The muted response of the job-finding rate slows the dynamic of the unemployment rate: While the response of unemployment peaks after five quarters under rational expectations, the peak is at quarter seven when agents use simple models.

The persistence bias in agents' expectations leads to additional persistence in vacancies, the job-finding rate, and the unemployment rate. The most economically-interesting decision in the DMP model is firms' vacancy-creation decision. Hiring a worker is a long-term investment. Therefore, it depends on forecasts of economic conditions by firms. Those forecasts are tied to the unemployment rate and show little response to changes in labor productivity, as can be seen in equation (21) and Figure 3.[41] The anchoring of expectations to unemployment dampens the initial response of vacancies to an increase in labor productivity.

The economy's response to a separation-rate shock is perhaps even more subtle. Figure 4 plots the impulse-response functions to a one-percent increase in the separation rate. Under rational expectations, the increase in the 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 unemployment rate in a DMP model with only separation-rate shocks.

Constraining agents to simple models turns this dynamic on its head. By equation (21), an increase in separations lowers agents' nowcast of the state of the economy both directly and

[40] I 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 $|p_u| + |p_a| + |p_s| = 1$.

[41] Since the scale of z_t is not identifiable either to agents or an outside observer, the scale of the nowcast $\hat{z}_t$ is intrinsically meaningless. However, Figures 3 and 4 use the same scale for $\hat{z}_t$, so the IRFs are comparable across the two figures.

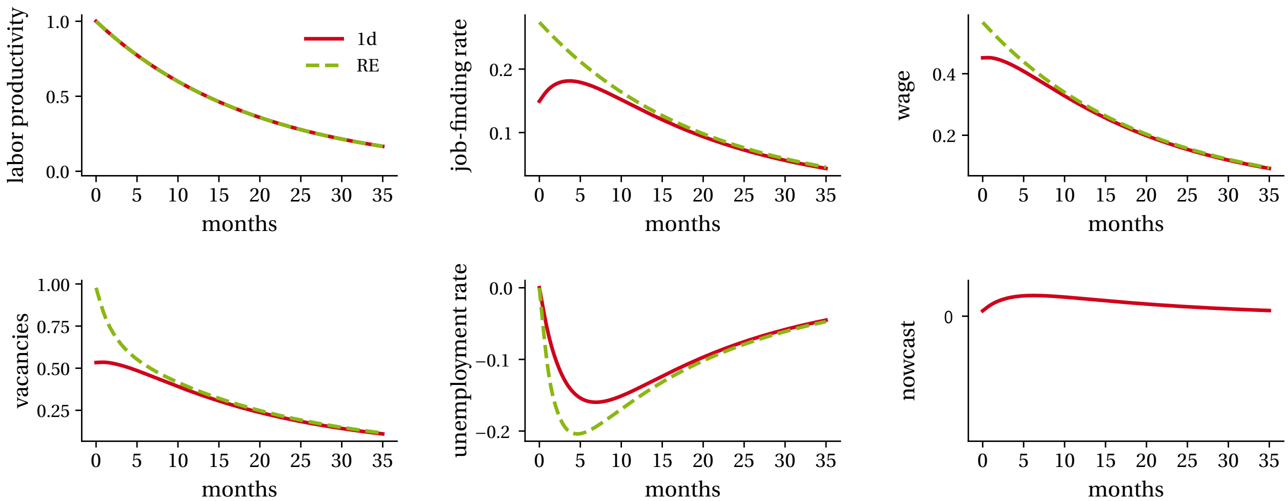

Figure 3. Impulse-response functions to a labor-productivity shock

Notes: This figure plots the impulse-response functions to a one-percent increase in labor productivity. The variable labeled "nowcast" is defined as agents' nowcast $\hat{z}_t$ of the subjective state z_t. Every variable, except for the nowcast, is measured in log changes from its steady-state value. The unit of measurement for agents' nowcast is arbitrary.

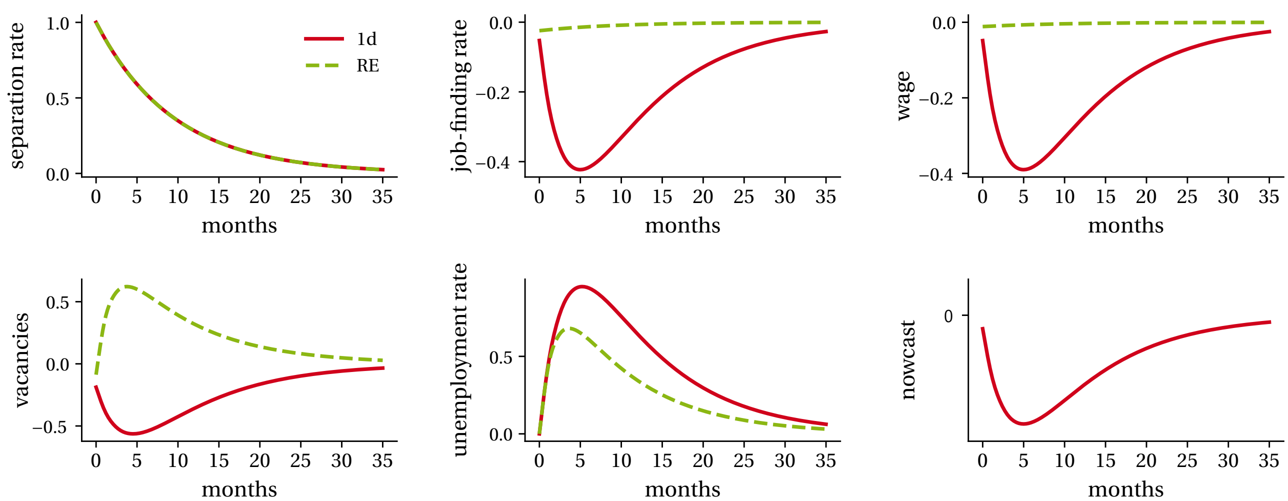

Figure 4. Impulse-response functions to a separation-rate shock

Notes: This figure plots the impulse-response functions to a one-percent increase in the separation rate. The variable labeled "nowcast" is defined as agents' nowcast $\hat{z}_t$ of the subjective state z_t. Every variable, except for the nowcast, is measured in log changes from its steady-state value. The unit of measurement for agents' nowcast is arbitrary.

indirectly through the resulting increase in unemployment. The deterioration in firms' nowcast lowers their expectations of returns to posting a vacancy. In the current calibration, this decrease is large enough 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 has a Keynesian flavor. The increase in separations and the resulting increase in unemployment make firms pessimistic. In response, they slow their recruiting activities, further exacerbating unemployment and worsening the economic outlook. The result is an inefficiently long and deep recession.

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. Although simple models do not help with the amplification of productivity shocks, they enable even the basic DMP model to exhibit significant amplification, propagation, and comovement in response to *separation shocks*.

8 Concluding Remarks

This paper suggests a novel approach to modeling bounded rationality, which is portable across different applications. I illustrated the use of the framework in three canonical macroeconomic models. In one application, I showed that constraining agents to use estimated simple models does not alter the divine coincidence property of the standard new-Keynesian model. However, it 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, one can easily incorporate simple models into modern heterogeneous-agent macro models. This allows one to examine how bounded rationality in the face of intertemporal complexity affects predictions of such models. This can be done because neither the additional degrees of freedom nor the computational burden of finding the equilibrium with simple 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).

Throughout the paper, I took dimension d of agents' model as a primitive parameter. This parameter can be identified using expectations data.[42] 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.

[42] See Molavi, Tahbaz-Salehi, and Vedolin (forthcoming) for a discussion of how this can be done in a closely related framework.

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 her 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 the weighted mean-squared error is in general not invariant to linear transformations of the vector of observable (unless if the weight matrix W is transformed accordingly). However, the following proposition establishes that mean-squared forecast-error minimization coincides with KLDR minimization under the appropriate choice of the weighting matrix W:

Proposition A.1. *Let θ denote a pseudo-true d-state model, and let $\hat{\Sigma}_y^\theta$ denote the implied subjective variance of y_{t+1} conditional on the agent's information at time t. If W is equal to the inverse of $\hat{\Sigma}_y^\theta$, then $\theta \in \arg\min_{\theta\in\Theta_d} MSE_W(\theta)$.*

The proof of the proposition is standard, and so, is omitted.

B Exponential Ergodicity

This appendix provides a set of sufficient conditions for a process to be exponentially ergodic and discusses the relationship between those conditions and the notion of full information.

Proposition B.1. *Consider a process $\mathbb{P}$ that can be represented as*

$$\begin{aligned} f_t &= F f_{t-1} + \epsilon_t \\ y_t &= H' f_t, \end{aligned} \tag{B.1}$$

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, $H \in \mathbb{R}^{m\times n}$, and 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 $\mathbb{P}$ is exponentially ergodic.

The assumption that the process has a representation of the form (B.1) is without loss of generality. The Wold representation theorem implies that any mean zero, covariance stationary, and purely non-deterministic process has a representation of this form (possibly with $m = \infty$). 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.[43] 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, for example, if F is diagonal or symmetric, but it is much weaker than symmetry.

The most substantial assumption of the proposition is the requirement that H is a rank-m matrix. This assumption can be seen as a full-information (or spanning) assumption: If the agent observes an observable of the form (B.1) 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 B.2. *Suppose the observable is one-dimensional, and the true process $\mathbb{P}$ can be represented as in* (B.1) *for some $f_t \in \mathbb{R}^m$, $\epsilon_t \sim \mathcal{N}(0, \Sigma)$, 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* (B.1) *is minimal and $m > 1$, then the s-period-ahead forecast of an agent who uses a pseudo-true one-state model θ is given by*

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

for some $a \in (-1, 1)$ and $\eta \in (0, 1)$.[44]

C 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 observable's law of motion depends on agents' choices. I consider a stylized general equilibrium (GE) economy in which observables are linear functions of 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

$$y_t^{\mathrm{GE}} = \tilde{H}' f_t + g x_t^{\mathrm{GE}}, \tag{C.1}$$

where $x_t \in \mathbb{R}$ is agents' time-t 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

[43] See Lemma D.5 of the Online Appendix and its proof for how this can be done.

[44] The representation in (B.1) of a process is *minimal* if there exists no representation for the process of the same form in which the dimension of f_t is strictly smaller.

agents' actions to the aggregate observable. Agents' best-response functions are given by

$$x_t^{\mathrm{GE}} = b' y_t^{\mathrm{GE}} + E_t\left[\sum_{s=1}^{\infty} \beta^s c' y_{t+s}^{\mathrm{GE}}\right]. \tag{C.2}$$

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

$$f_t = F f_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \Sigma), \tag{C.3}$$

where $F = \mathrm{diag}(\alpha_1, \ldots, \alpha_m)$ and $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2)$. Equations (C.1)–(C.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

$$y_t^{\mathrm{PE}} = H' f_t, \tag{C.4}$$

$$x_t^{\mathrm{PE}} = b' y_t^{\mathrm{PE}} + E_t\left[\sum_{s=1}^{\infty} \beta^s c' y_{t+s}^{\mathrm{PE}}\right], \tag{C.5}$$

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

Proposition C.1. *Consider the general equilibrium economy* (C.1)–(C.3) *and the partial equilibrium economy* (C.3)–(C.5)*, and suppose that, in each economy, agents use pseudo-true Markovian 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 1) 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

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.

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.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, 113, 1–45.

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.

Del Negro, Marco, Marc P. Giannoni, and Christina Patterson (2023), "The Forward Guidance Puzzle." *Journal of Political Economy Macroeconomics*, 1, 43–79.

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.

Farhi, Emmanuel and Iván Werning (2019), "Monetary Policy, Bounded Rationality, and Incomplete Markets." *American Economic Review*, 109, 3887–3928.

Faurre, 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.

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 Iourii 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.

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.

Molavi, Pooya (2019), "Macroeconomics with Learning and Misspecification: A General Theory and Applications." Working paper.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Andrea Vedolin (forthcoming), "Model Complexity, Expectations, and Asset Prices." *Review of Economic Studies*.

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.

Preston, Bruce (2005), "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter." *International Journal of Central Banking*.

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.

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.

Smets, Frank and Rafael Wouters (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97, 586–606.

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

D Proofs

Proof of Theorem 1

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 $\mathrm{KLDR}(\theta) = +\infty$. In what follows, I assume that P^θ is supported on $\mathcal{W}$.

Note that minimizing the KLDR over the set Θ_d of d-state models is equivalent to minimizing the KLDR over the set $\Theta_0^m \cup \Theta_1^m \cup \cdots \cup \Theta_d^m$, where Θ_k^m denotes the set of models whose minimal realization requires k state variables. Therefore, in the proofs, I assume without loss of generality that the d-state model θ is minimal, i.e., that there exists no d'-state model with $d' < d$ that is observationally equivalent to θ.

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

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

Furthermore, by the chain rule,

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

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

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

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

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

where the second equality uses the stationarity of P^θ, and the fact that $\log f^\theta(y_{t+1}|y_t, \ldots)$ is well defined is a consequence of the assumption that A is convergent and Q is positive definite for any $\theta = (A, B, Q, R)$. The above display implies that the Cesàro sum also converges:

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

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 $\{y_t, y_{t-1}, \ldots\}$.

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:

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

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

$$K \equiv A\hat{\Sigma}_z B\left(B'\hat{\Sigma}_z B + R\right)^\dagger, \tag{D.2}$$

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:[45] [46]

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

Solving equation (D.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 subjective conditional variance of y_{t+1} can be expressed in terms of the subjective conditional variance of z_{t+1}:

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

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

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

The Kullback–Leibler divergence rate is thus equal to

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

[45] Note 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.

[46] The assumptions that the d-state model θ is minimal and Q is positive definite imply that the Riccati equation has a unique positive semidefinite solution and that $A - KB'$ is a convergent matrix.

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.[47] 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.

Proof of Theorem 1. 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 $\tilde{y}_t \in \mathbb{R}^{\tilde{n}}$, and let $\widetilde{\text{KLDR}}(\tilde{\theta})$ denote the KLDR of model $\tilde{\theta} \in \widetilde{\Theta}_d$ from the true process $\widetilde{\mathbb{P}} \equiv T(\mathbb{P})$.

Let $\theta \in \Theta_d$ denote an arbitrary pseudo-true d-state model when the true process is $\mathbb{P}$ and $\tilde{\theta} \in \widetilde{\Theta}_d$ denote an arbitrary pseudo-true d-state model when the true process is $\widetilde{\mathbb{P}}$. I first show that $T(P^\theta)$ and $P^{\tilde{\theta}}$ are both supported on $\widetilde{\mathcal{W}}$. Note that there always exists a d-state model for which the KLDR is finite—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, P^θ is supported on $\mathcal{W}$, and so, $T(P^\theta)$ is supported on $\widetilde{\mathcal{W}}$. On the other hand, since the true distribution $\mathbb{P}$ is supported on $\mathcal{W}$, the transformed distribution $\widetilde{\mathbb{P}}$ is supported on $\widetilde{\mathcal{W}}$. Consequently, by the above argument, $P^{\tilde{\theta}}$ 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 $\widetilde{\hat{\Sigma}}_z \equiv \text{Var}_t^{T(\theta)}(z_{t+1})$ denote the corresponding conditional variance under model $T(\theta)$. Matrices $\hat{\Sigma}_z$ and $\widetilde{\hat{\Sigma}}_z$ solve the following Riccati equations:

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
\widetilde{\text{KLDR}}(T(\theta)) = &-\frac{1}{2}\log\det{}^*\left(T^{\dagger\prime}\hat{\Sigma}_y^{\dagger}T^{\dagger}\right) + \frac{n}{2}\log(2\pi) + \frac{1}{2}\text{tr}\left(T^{\dagger\prime}\hat{\Sigma}_y^{\dagger}T^{\dagger}T\Gamma_0 T'\right) \\
&-\frac{1}{2}\sum_{\tau=1}^{\infty}\text{tr}\left(T^{\dagger\prime}\hat{\Sigma}_y^{\dagger}T^{\dagger}T\Phi_\tau T^{\dagger}T\Gamma_\tau' T'\right) - \frac{1}{2}\sum_{\tau=1}^{\infty}\text{tr}\left(T^{\dagger\prime}\hat{\Sigma}_y^{\dagger}T^{\dagger}T\Gamma_\tau T'T^{\dagger\prime}\Phi_\tau' T'\right) \\
&+\frac{1}{2}\sum_{s=1}^{\infty}\sum_{\tau=1}^{\infty}\text{tr}\left(T^{\dagger\prime}\hat{\Sigma}_y^{\dagger}T^{\dagger}T\Phi_s T^{\dagger}T\Gamma_{\tau-s}T'T^{\dagger\prime}\Phi_\tau' T'\right) + \text{constant}.
\end{aligned}$$

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 $\tilde{\theta} = (\tilde{A}, \tilde{B}, \tilde{Q}, \tilde{R}) \in \widetilde{\Theta}_d$, define $T^{-1}(\tilde{\theta}) \equiv (\tilde{A}, \tilde{B}T^{\dagger\prime}, \tilde{Q}, T^{\dagger}\tilde{R}T^{\dagger\prime}) \in \Theta_d$. 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 $\widetilde{\Theta}_d$: Any model $\theta \in \Theta_d$ can be identified with a model $T(\theta) \in \widetilde{\Theta}_d$ such that the KLDR of P^{θ} from the process $\mathbb{P}$ is equal to the KLDR of $P^{T(\theta)}$ from $T(\mathbb{P})$, and any model $\tilde{\theta} \in \widetilde{\Theta}_d$ 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 $T(\mathbb{P})$. 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 $T(\mathbb{P})$.

It only remains to show that $P^{T(\theta)} = T(P^{\theta})$ for any model $\theta \in \Theta_d$. Since $P^{T(\theta)}$ and $T(P^{\theta})$ 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 $T(P^{\theta})$. □

Proof of Theorem 2

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

Lemma D.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 $R = N^{-1\prime}(I - DD')N^{-1}$, where (M, D, N) is a tuple that minimizes*

$$KLDR(\tilde{M}, \tilde{D}, \tilde{N}) \equiv -\frac{1}{2}\log det(\tilde{N}\tilde{N}') + \frac{1}{2}\text{tr}(\tilde{N}'\Gamma_0\tilde{N}) - \sum_{\tau=1}^{\infty}\text{tr}\left(\left(\tilde{M}(I - \tilde{D}'\tilde{D})\right)^{\tau-1}\tilde{M}\tilde{D}'\tilde{N}'\Gamma_\tau'\tilde{N}\tilde{D}\right)$$

$$+\frac{1}{2}\sum_{s=1}^{\infty}\sum_{\tau=1}^{\infty}\operatorname{tr}\left(\tilde{D}\left(\tilde{M}\left(I-\tilde{D}'\tilde{D}\right)\right)^{s-1}\tilde{M}\tilde{D}'\tilde{N}'\Gamma_{\tau-s}\tilde{N}\tilde{D}\tilde{M}'\left(\left(I-\tilde{D}'\tilde{D}\right)\tilde{M}'\right)^{\tau-1}\tilde{D}'\right) \tag{D.9}$$

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 (D.3) is invertible. On the other hand, since Γ_0 is invertible, I can restrict attention to subjective models for which $\hat{\Sigma}_y$ is non-singular.[48] The pseudo-inverses and pseudo-determinants in equations (D.3) and (D.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:

$$\hat{\Sigma}_y^{\frac{-1}{2}}B'\hat{\Sigma}_z^{\frac{1}{2}}=UDV', \tag{D.10}$$

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

$$VD'DV'=\hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_zB+R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}. \tag{D.11}$$

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,

$$A=\hat{\Sigma}_z^{\frac{1}{2}}VMV^{-1}\hat{\Sigma}_z^{\frac{-1}{2}}, \tag{D.12}$$

$$B=\hat{\Sigma}_z^{\frac{-1}{2}}VD'U'\hat{\Sigma}_y^{\frac{1}{2}}, \tag{D.13}$$

$$K=\hat{\Sigma}_z^{\frac{1}{2}}VMD'U'\hat{\Sigma}_y^{\frac{-1}{2}}, \tag{D.14}$$

and so

$$KB'=\hat{\Sigma}_z^{\frac{1}{2}}VMD'DV'\hat{\Sigma}_z^{\frac{-1}{2}},$$
$$\Phi_\tau=\hat{\Sigma}_y^{\frac{1}{2}}UD\left(M\left(I-D'D\right)\right)^{\tau-1}MD'U'\hat{\Sigma}_y^{\frac{-1}{2}}.$$

Note that since A is a convergent matrix, so is M. Substituting in (D.3) for A from equation (D.12) and for B from (D.13), I get

$$\begin{aligned}Q&=\hat{\Sigma}_z-A\left(\hat{\Sigma}_z-\hat{\Sigma}_zB\left(B'\hat{\Sigma}_zB+R\right)^{-1}B'\hat{\Sigma}_z\right)A'\\&=\hat{\Sigma}_z-\hat{\Sigma}_z^{\frac{1}{2}}VMV^{-1}\hat{\Sigma}_z^{\frac{-1}{2}}\left(\hat{\Sigma}_z-\hat{\Sigma}_z^{\frac{1}{2}}VD'DV\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{-1}{2}}VM'V^{-1}\hat{\Sigma}_z^{\frac{1}{2}}\\&=\hat{\Sigma}_z-\hat{\Sigma}_z^{\frac{1}{2}}VM\left(I-D'D\right)M'V^{-1}\hat{\Sigma}_z^{\frac{1}{2}}.\end{aligned} \tag{D.15}$$

[48] 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 1, the restriction to true processes with invertible variance-covariance matrices is without loss of generality.

Therefore, since Q is positive definite, the eigenvalues of $VM\,(I-D'D)\,M'V^{-1}$ 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

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

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

$$\operatorname{tr}\left(\hat{\Sigma}_y^{-1}\Gamma_0\right) = \operatorname{tr}\left(\hat{\Sigma}_y^{\frac{-1}{2}}\Gamma_0\hat{\Sigma}_y^{\frac{-1}{2}}\right) = \operatorname{tr}\left(UN'\Gamma_0NU'\right) = \operatorname{tr}\left(N'\Gamma_0N\right).$$

On the other hand,

$$\begin{aligned}\operatorname{tr}\left(\hat{\Sigma}_y^{-1}\Phi_\tau\Gamma_\tau'\right) &= \operatorname{tr}\left(\hat{\Sigma}_y^{\frac{-1}{2}} UD\left(M\,(I-D'D)\right)^{\tau-1} MD'U'\hat{\Sigma}_y^{\frac{-1}{2}}\Gamma_\tau'\right)\\ &= \operatorname{tr}\left(\left(M\,(I-D'D)\right)^{\tau-1} MD'N'\Gamma_\tau'ND\right),\end{aligned}$$

and

$$\operatorname{tr}\left(\hat{\Sigma}_y^{-1}\Phi_s\Gamma_{\tau-s}\Phi_\tau'\right) = \operatorname{tr}\left(D\left(M\,(I-D'D)\right)^{s-1} MD'N'\Gamma_{\tau-s}NDM'\left((I-D'D)\,M'\right)^{\tau-1}D'\right).$$

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

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

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 (D.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 (D.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

$$K = MD'N'. \tag{D.16}$$

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

$$E_t^\theta[y_{t+s}] = N'^{-1}DM^{s-1}\sum_{\tau=0}^{\infty}\left(M\left(I - D'D\right)\right)^\tau MD'N'y_{t-\tau}. \tag{D.17}$$

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

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

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

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

Solving the above equation forward, I get

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

Therefore,

$$\mathrm{Var}^\theta(y) = B'\sum_{\tau=0}^{\infty}A^\tau QA'^\tau B + R = N^{-1'}\left(I + \sum_{\tau=1}^{\infty}DM^\tau D'DM'^\tau D'\right)N^{-1}. \tag{D.18}$$

□

I can now establish Theorem 2.

Proof of Theorem 2. Let M, D, and N be as in Lemma D.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 (D.9), can be written (with slight abuse of notation) as a function of a, η, and S:

$$\text{KLDR}(a,\eta,S) = -\frac{1}{2}\log\det(SS') + \frac{1}{2}\operatorname{tr}(S'S) - \frac{1}{2}e_1'S'\Omega(a,\eta)Se_1 + \text{constant},$$

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:

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

Therefore,

$$\Omega(a,\eta) = -\frac{a^2(1-\eta)^2}{1-a^2\eta^2}I + \frac{(1-\eta)(1-a^2\eta)}{1-a^2\eta^2}\sum_{\tau=1}^{\infty}a^\tau\eta^{\tau-1}\Gamma_0^{\frac{-1}{2}}(\Gamma_\tau+\Gamma_\tau')\Gamma_0^{\frac{-1}{2}}. \tag{D.19}$$

By Lemma D.1, minimizing the KLDR with respect to A, B, Q, and R is equivalent to minimizing 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 KLDR(a,η,S) with respect to a, η, and S.

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

$$S^{-1} = S' - e_1e_1'S'\Omega(a,\eta),$$

which implies that

$$S'S - e_1e_1'S'\Omega(a,\eta)S = I. \tag{D.20}$$

Therefore, for any solution to the problem of minimizing KLDR(a,η,S),

$$n = \operatorname{tr}(I) = \operatorname{tr}(S'S) - \operatorname{tr}\left(e_1e_1'S'\Omega(a,\eta)S\right) = \operatorname{tr}(S'S) - e_1'S'\Omega(a,\eta)Se_1.$$

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

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

where

$$S(a,\eta) \in \arg\min_{S} \ -\frac{1}{2}\log\det\left(SS'\right) + \frac{1}{2}\operatorname{tr}\left(S'S\right) - \frac{1}{2}e_1'S'\Omega(a,\eta)Se_1. \tag{D.21}$$

I proceed by first characterizing $S(a,\eta)$. Note that the necessary first-order optimality conditions for problem (D.21) are given by matrix equation (D.20).

Claim D.1. *For any matrix S that solves equation* (D.20)*, the necessary first-order optimality condition for problem* (D.21)*,*

(i) $Se_1 = \dfrac{1}{\sqrt{1-\lambda}}u$,

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

(iii) $SS' = I + \frac{\lambda}{1-\lambda}uu'$,

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 (D.20) in general has multiple solutions, with each solution corresponding to a local extremum of problem (D.21). The global optimum of problem (D.21) is given by the solution to equation (D.20) that results in the largest value for $\det(SS')$. But by part (iii) of Claim D.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 D.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 D.1 is thus given by[49]

$$\begin{aligned}
A &= a,\\
B &= \sqrt{1-\eta}e_1'S^{-1}\Gamma_0^{\frac{1}{2}},\\
Q &= 1 - a^2\eta,\\
R &= \Gamma_0^{\frac{1}{2}}S^{-1\prime}\left(I - (1-\eta)e_1e_1'\right)S^{-1}\Gamma_0^{\frac{1}{2}}.
\end{aligned}$$

[49] For 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 Lemma D.3.

By Claim D.1 and the argument above,

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

Thus,

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

and

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

Finally, note that $M = a$, $D = \sqrt{1-\eta}\, e_1$, and $N = \Gamma_0^{\frac{-1}{2}} S$. Therefore, by equation (D.17), the subjective forecasts are given by

$$
E_t^{\theta}[y_{t+s}] = a^s (1-\eta) \Gamma_0^{\frac{1}{2}} S'^{-1} e_1 e_1' S' \Gamma_0^{\frac{-1}{2}} \sum_{\tau=0}^{\infty} a^{\tau} \eta^{\tau} y_{t-\tau}. \tag{D.22}
$$

Using Claim D.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. □

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

$$
S'S - e_1 e_1' S' \Omega S = I. \tag{D.23}
$$

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

$$
Se_1 - \Omega S e_1 = S'^{-1} e_1. \tag{D.24}
$$

On the other hand, multiplying equation (D.23) from left by S and from right by S^{-1}, I get

$$
SS' = I + S e_1 e_1' S' \Omega. \tag{D.25}
$$

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 Se_1, I get

$$
S'^{-1} e_1 = \frac{1}{1 + e_1' S' \Omega S e_1} S e_1. \tag{D.26}
$$

Substituting for $S'^{-1}e_1$ from the above equation in (D.24) and rearranging the terms, I get

$$\Omega S e_1 = \frac{e_1' S' \Omega S e_1}{1 + e_1' S' \Omega S e_1} S e_1. \tag{D.27}$$

That is, Se_1 is an eigenvector of Ω. Let λ denote the corresponding eigenvalue and let $u = Se_1/\sqrt{e_1'S'Se_1}$. Then equation (D.27) 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 (D.26) then implies that

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

and equation (D.25) implies that

$$SS' = I + \frac{\lambda}{1-\lambda} uu'.$$

If $\lambda = 0$, then equation (D.24) implies that $Se_1 = S'^{-1}e_1$, and so, Se_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, $Se_1 = S'^{-1}e_1 = u$. On the other hand, equation (D.25) implies that $SS' = I$. This completes the proof of the claim. □

Proof of Theorem 3

I first prove two useful lemmas:

Lemma D.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. □

Lemma D.3. *If $\mathbb{P}$ is purely non-deterministic and stationary ergodic, then so is P^θ for any pseudo-true one-state model θ.*

Proof. Define

$$C(a,\eta) \equiv \sum_{\tau=1}^{\infty} a^\tau \eta^{\tau-1} C_\tau. \tag{D.28}$$

Then

$$\lambda_{\max}(\Omega(a,\eta)) = -\frac{a^2(1-\eta)^2}{1-a^2\eta^2} + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2}\lambda_{\max}(C(a,\eta)), \tag{D.29}$$

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 $\lambda_{\max}(\Omega(a,\eta))$.[50] The partial derivatives of $\lambda_{\max}(\Omega(a,\eta))$ with respect to a and η are given by

$$\begin{aligned}
\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial a} &= \frac{-2a(1-\eta)^2}{\left(1-a^2\eta^2\right)^2} - \frac{4a\eta(1-\eta)^2}{\left(1-a^2\eta^2\right)^2}\lambda_{\max}(C) \\
&\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2} u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C), && \text{(D.30)} \\
\frac{\partial \lambda_{\max}(\Omega(a,\eta))}{\partial \eta} &= \frac{2a^2(1-\eta)(1-a^2\eta)}{\left(1-a^2\eta^2\right)^2} - \frac{2\left(1+a^4\eta^2+a^2(1-4\eta+\eta^2)\right)}{\left(1-a^2\eta^2\right)^2}\lambda_{\max}(C) \\
&\quad + \frac{2(1-\eta)(1-a^2\eta)}{1-a^2\eta^2} u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C), && \text{(D.31)}
\end{aligned}$$

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

$$\eta u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C) + \lambda_{\max}(C) = a u'_{\max}(C)\frac{\partial C}{\partial a}u_{\max}(C). \tag{D.32}$$

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

[50] The 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.

in the proof of Theorem 2, the pseudo-true one-state model is i.i.d., and $A = a^*$ can be chosen arbitrarily to satisfy $|a^*| < 1$.[51]

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

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

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

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

where in the second equality I am using identity (D.32). Therefore,

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

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 second inequality is by Lemma D.2. Therefore,

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

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 that the one-state model with $a = a^*$ and $\eta = \eta^*$ is purely non-deterministic and stationary ergodic. □

I can now prove the theorem.

[51] The pseudo-true one-state model then has a zero-state minimal representation.

Proof of Theorem 3. Setting $M = a$, $D = \sqrt{1-\eta}e_1$, and $N = \Gamma_0^{\frac{-1}{2}} S$ in equation (D.18), I get

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

where a, η, $\lambda = \lambda_{\max}(\Omega(a,\eta))$, and u are as in Theorem 2. Substituting for $\lambda_{\max}(\Omega(a,\eta))$ from equation (D.29) in the above equation, I get

$$\mathrm{Var}^\theta(y_t) = \Gamma_0^{\frac{1}{2}} \left[I + \frac{2(1-\eta)(1-a^2\eta)}{(1-a^2)(1-a^2\eta^2)} \left(a^2(1-\eta) - (1-2a^2\eta + a^2\eta^2)\lambda_{\max}(C) \right) uu' \right] \Gamma_0^{\frac{1}{2}}. \tag{D.33}$$

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 (D.33) is equal to Γ_0.

Next suppose $\eta^* < 1$. By the argument in the proof of Lemma D.3, 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 (D.30) and multiplying both sides of the equation by a^*, I get, using (D.32),

$$\begin{aligned} &\frac{2a^{*2}(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2} + \frac{4a^{*2}\eta^*(1-\eta^*)^2}{\left(1-a^{*2}\eta^{*2}\right)^2}\lambda_{\max}(C) \\ &\quad= \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}}\lambda_{\max}(C) + \frac{2(1-\eta^*)(1-a^{*2}\eta^*)}{1-a^{*2}\eta^{*2}}\eta^* u'_{\max}(C)\frac{\partial C}{\partial \eta}u_{\max}(C). \end{aligned} \tag{D.34}$$

Setting $\eta^* = 0$ in the above equation, I get $a^{*2} = \lambda_{\max}(C)$. Setting $a^{*2} = \lambda_{\max}(C)$ in equation (D.33) then establishes that $\mathrm{Var}^\theta(y_t) = \Gamma_0$ 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 (D.31), multiplying it by η^*, solving for $\eta^* u'_{\max}(C)\frac{\partial C}{\partial\eta}u_{\max}(C)$, and substituting in equation (D.34), I get

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

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 (D.33) implies that $\mathrm{Var}^\theta(y_t) = \Gamma_0$ and finishes the proof of the theorem. □

Proof of Theorem 4

Let λ denote the eigenvalue of C_1 largest in magnitude.[52] 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,

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

The above expression is maximized by setting $(1-\eta)a = 0$. Therefore, by Theorem 2, 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

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

where in the second equality I am using the fact that $\rho(C_\tau) < 1$, established in Lemma D.2. Function $\overline{f}(a,\eta)$ has two maximizers given by $(\overline{a}^*, \overline{\eta}^*) = (-\rho(C_1), 0)$ and $(\overline{a}^*, \overline{\eta}^*) = (\rho(C_1), 0)$ with the maximum given by $\overline{f}^* = \rho(C_1)^2$. I establish the theorem by showing that $\lambda_{\max}(\Omega(a,\eta)) \leq \overline{f}(a,\eta)$ for all a and η, $\lambda_{\max}(\Omega(\lambda,0)) = \overline{f}(\lambda,0) = \overline{f}^*$, and $\lambda_{\max}(\Omega(-\lambda,0)) \leq \overline{f}(-\lambda,0) = \overline{f}^*$ 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,

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

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

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

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

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

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

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

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

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

[52] The 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 $|\lambda| = |\lambda'| = \rho(C_1)$.

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,

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

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

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

Thus,

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

and

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

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

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

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 2, 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. □

Proof of Proposition 2

I first state and prove a lemma, which is used in the proof of the proposition.

Lemma D.4. *Any Markovian model θ has a representation as in Lemma D.1 for which $D'D = I$.*

Proof. Fix a Markovian model θ, and let M, D, and N be as in Lemma D.1. By (D.17), the s-step-ahead forecast under model θ is given by

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

Since θ is Markovian and N is invertible, $D\left(M\left(I-D'D\right)\right)^{\tau} MD' = \mathbf{0}$ for all $\tau \geq 1$. As the first step of the proof, I use this identity and an inductive argument to show that $DM^sD' = (DMD')^s$ for all $s \geq 2$. The following equation establishes the induction base:

$$\mathbf{0} = D\left(M\left(I-D'D\right)\right) MD' = DM^2D' - DMD'DMD'.$$

As the induction hypothesis, suppose $DM^sD' = (DMD')^s$ for some $s \geq 2$. Note that

$$\begin{aligned} D\left(M\left(I-D'D\right)\right)^{s} MD' &= D\left(M\left(I-D'D\right)\right)^{s-1} M(I-D'D)MD' \\ &= D\left(M\left(I-D'D\right)\right)^{s-1} M^2D' - D\left(M\left(I-D'D\right)\right)^{s-1} MD'DMD' \\ &= D\left(M\left(I-D'D\right)\right)^{s-1} M^2D', \end{aligned}$$

where the last equality follows the fact that $D\left(M\left(I-D'D\right)\right)^{s-1} MD' = \mathbf{0}$ for any $s \geq 2$. By a similar argument,

$$D\left(M\left(I-D'D\right)\right)^{s} MD' = D\left(M\left(I-D'D\right)\right)^{s-2} M^3D' = \cdots = D\left(M\left(I-D'D\right)\right) M^sD'.$$

Therefore,

$$D\left(M\left(I-D'D\right)\right)^{s} MD' = DM^{s+1}D' - DMD'DM^sD' = DM^{s+1}D' - (DMD')^{s+1},$$

where the last equality follows the induction hypothesis. The assumption that $D\left(M\left(I-D'D\right)\right)^{s} MD' = \mathbf{0}$ then proves the induction step.

I next find a model $\tilde{\theta}$, represented by matrices $\tilde{M}$, $\tilde{D}$, and $\tilde{N}$, that is observationally equivalent to θ and for which $\tilde{D}'\tilde{D} = I$. Since $D \in \mathbb{R}^{n\times d}$ is a rectangular diagonal matrix and $d \leq n$,

$$DMD' = \begin{pmatrix} M_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

for some $d \times d$ matrix M_1. Let $\tilde{M} = M_1$, $\tilde{D} \in \mathbb{R}^{n\times d}$ be the rectangular diagonal matrix with its diagonal elements equal to one, and $\tilde{N} = N$. Then $\tilde{D}'\tilde{D} = I$. Furthermore,

$$DM^sD' = (DMD')^s = \begin{pmatrix} M_1^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \tilde{D}\tilde{M}^s\tilde{D}'.$$

By equation (D.17), the forecasts are identical under the two models:

$$E_t^{\theta}[y_{t+s}] = N'^{-1}DM^sD'N'y_t = \tilde{N}'^{-1}\tilde{D}\tilde{M}^s\tilde{D}'\tilde{N}'y_t = E_t^{\tilde{\theta}}[y_{t+s}].$$

By equation (D.18), the unconditional variance of the observable is also identical under the two models:

$$\mathrm{Var}^{\theta}(y) = N'^{-1}\left(I + \sum_{\tau=1}^{\infty} DM^{\tau}D'DM'^{\tau}D'\right)N^{-1} = \tilde{N}'^{-1}\left(I + \sum_{\tau=1}^{\infty} \tilde{D}\tilde{M}^{\tau}\tilde{D}'\tilde{D}\tilde{M}'^{\tau}\tilde{D}'\right)\tilde{N}^{-1} = \mathrm{Var}^{\tilde{\theta}}(y).$$

On the other hand,

$$E^{\theta}[y_{t+s}y_t'] = E^{\theta}[E_t^{\theta}[y_{t+s}]y_t'] = N'^{-1}DM^sD'N'E^{\theta}[y_ty_t'] = N'^{-1}DM^sD'N'\mathrm{Var}^{\theta}(y),$$

and similarly for $E^{\tilde{\theta}}[y_{t+s}y_t']$. Therefore, $E^{\theta}[y_{t+s}y_t'] = E^{\tilde{\theta}}[y_{t+s}y_t']$ for all s; that is, P^{θ} and $P^{\tilde{\theta}}$ also have identical autocovariance matrices at all lags. This conclusion, together with the fact that P^{θ} and $P^{\tilde{\theta}}$ are both zero-mean Gaussian distributions, implies that they are observationally equivalent. □

I can now prove the proposition.

Proof of Proposition 2. By equation (D.4),

$$\mathrm{Var}_t^{\theta}(y_{t+1}) = B'\hat{\Sigma}_z B + R,$$

where $\hat{\Sigma}_z$ solves the algebraic Riccati equation (D.3). The equation can be written as

$$\hat{\Sigma}_z = A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{1}{2}}A' + Q. \tag{D.35}$$

Since R is a positive semidefinite matrix, so is $I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}$. Therefore,

$$\hat{\Sigma}_z \geq Q,$$

and so,

$$\mathrm{Var}_t^{\theta}(y_{t+1}) \geq B'QB + R.$$

On the other hand,

$$\mathrm{Var}^{\theta}(y_{t+1}|z_t) = B'\mathrm{Var}^{\theta}(z_{t+1}|z_t)B + R = B'QB + R,$$

where I am using the assumption that w_t is i.i.d. $\mathcal{N}(0,Q)$, v_t is i.i.d. $\mathcal{N}(0,R)$, and w_t and v_t are independent. This proves the first part of the proposition.

To prove the second part, first assume that $\mathrm{Var}_t^{\theta}(y_{t+1}) = B'QB + R$. Together with equation (D.35), this implies that

$$B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{1}{2}}A'B = \mathbf{0}.$$

Since $\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)$ is a symmetric positive semidefinite matrix, the above equation implies that

$$B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right) = \mathbf{0}.$$

On the other hand, by equation (D.5), the one-step-ahead forecast under model θ is given by

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

Substituting for K from equation (D.2), I get

$$B'(A - KB') = B'\left(A - A\hat{\Sigma}_z B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\right) = B'A\hat{\Sigma}_z^{\frac{1}{2}}\left(I - \hat{\Sigma}_z^{\frac{1}{2}}B\left(B'\hat{\Sigma}_z B + R\right)^{-1}B'\hat{\Sigma}_z^{\frac{1}{2}}\right)\hat{\Sigma}_z^{\frac{-1}{2}} = \mathbf{0}.$$

Therefore,

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

On the other hand, $\mathrm{Var}_t^{\theta}(y_{t+1}) = B'\hat{\Sigma}_z B + R$. Under model θ, the mean and variance of y_{t+1} conditional on $\{y_\tau\}_{\tau\leq t}$ are both independent of $\{y_\tau\}_{\tau<t}$. Furthermore, P^{θ} is Gaussian. Therefore, it is Markovian.

Next, suppose P^{θ} is Markovian. Then by Lemma D.4, model θ has a representation as in Lemma D.1 for which $D'D = I$. By equation (D.11), then

$$\hat{\Sigma}_z^{\frac{1}{2}} B \left(B'\hat{\Sigma}_z B + R\right)^{-1} B'\hat{\Sigma}_z^{\frac{1}{2}} = VD'DV' = I,$$

where the second equality follows the facts that $D'D = I$ and V is orthogonal. Substituting in equation (D.35), I get $\hat{\Sigma}_z = Q$. Therefore,

$$\mathrm{Var}_t^{\theta}(y_{t+1}) = B'\hat{\Sigma}_z B + R = B'QB + R.$$

This completes the proof of the proposition. □

Proof of Theorem 5

By Lemma D.1, the agent's model can be represented in terms of matrices M, D, and N. Since the agent is restricted to the set of Markovian models, by Lemma D.4, I can set $D = (I\ \mathbf{0})'$. Let $S \equiv \Gamma_0^{\frac{1}{2}} N$ and $\Gamma \equiv \Gamma_0^{\frac{-1}{2}} \Gamma_1 \Gamma_0^{\frac{-1}{2}}$. The expression for the KLDR in (D.9) then simplifies to

$$\mathrm{KLDR}(M, S, D) = -\frac{1}{2} \log\det\left(SS'\right) + \frac{1}{2}\,\mathrm{tr}\left(S'S\right) - \mathrm{tr}\left(MD'S'\Gamma SD\right) + \frac{1}{2}\,\mathrm{tr}\left(MD'S'SDM'\right) + \text{constant}.$$

Write $S = (S_1\ S_2)$, where $S_1 \in \mathbb{R}^{n\times d}$ and $S_2 \in \mathbb{R}^{n\times(n-d)}$. The above expression can then be written as

$$-\frac{1}{2}\log\det\left(S_1S_1' + S_2S_2'\right) + \frac{1}{2}\,\mathrm{tr}\left(S_1'S_1\right) + \frac{1}{2}\,\mathrm{tr}\left(S_2'S_2\right) - \mathrm{tr}\left(MS_1'\Gamma S_1\right) + \frac{1}{2}\,\mathrm{tr}\left(MS_1'S_1M'\right) + \text{constant}.$$

I next optimize the above expression with respect to M, S_1, and S_2. The first-order optimality condition with respect to S_2 is given by

$$\left(S_1S_1' + S_2S_2'\right)^{-1} S_2 = S_2.$$

The above equation can be written as

$$S_1S_1'S_2 + S_2(S_2'S_2 - I) = \mathbf{0}.$$

Let b_0 be an arbitrary vector in $\mathbb{R}^{n-d}$, $b_1 \equiv S_1'S_2b_0 \in \mathbb{R}^d$, and $b_2 \equiv (S_2'S_2 - I)b_0 \in \mathbb{R}^{n-d}$. The above equation then implies that

$$0 = \left(S_1S_1'S_2 + S_2(S_2'S_2 - I)\right) b_0 = S_1b_1 + S_2b_2 = Sb,$$

where $b \equiv \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^n$. Since S is an invertible matrix, it must be that $b = 0$. Therefore, $b_1 = 0$ and $b_2 = 0$. Since b_0 was arbitrary, $S_2'S_2 = I$ and $S_1'S_2 = \mathbf{0}$. On the other hand,

$$\log\det\left(S_1S_1' + S_2S_2'\right) = \log\det(SS') = \log\det(S'S) = \log\det\begin{pmatrix} S_1'S_1 & S_1'S_2 \\ S_2'S_1 & S_2'S_2 \end{pmatrix}.$$

Therefore,

$$\log\det\left(S_1S_1' + S_2S_2'\right) = \log\det\begin{pmatrix} S_1'S_1 & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix} = \log\det(S_1'S_1).$$

The KLDR can thus be written only as a function of M and S_1 as

$$\text{KLDR}(M, S_1) = -\frac{1}{2}\log\det\left(S_1'S_1\right) + \frac{1}{2}\operatorname{tr}\left(S_1'S_1\right) - \operatorname{tr}\left(MS_1'\Gamma S_1\right) + \frac{1}{2}\operatorname{tr}\left(MS_1'S_1M'\right) + \text{constant}. \tag{D.36}$$

The first-order optimality conditions with respect to M and S_1 are then given by

$$-S_1'\Gamma'S_1 + MS_1'S_1 = 0, \tag{D.37}$$

$$-S_1^{\dagger\prime} + S_1 - \Gamma S_1 M - \Gamma' S_1 M' + S_1 M'M = 0. \tag{D.38}$$

Since $S_1'S_1$ is invertible, (D.37) can be solve for M to get $M = S_1'\Gamma'S_1(S_1'S_1)^{-1}$. Substituting in (D.38), I get

$$S_1(S_1'S_1)^{-1} = S_1 - \Gamma S_1 S_1'\Gamma' S_1(S_1'S_1)^{-1} - \Gamma' S_1(S_1'S_1)^{-1}S_1'\Gamma S_1 + S_1(S_1'S_1)^{-1}S_1'\Gamma S_1 S_1'\Gamma' S_1(S_1'S_1)^{-1}, \tag{D.39}$$

where I am using the fact that $S_1^{\dagger} = (S_1'S_1)^{-1}S_1'$. Next consider the singular-value decomposition of S_1:

$$S_1 = U\Sigma V', \tag{D.40}$$

where $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{d\times d}$ are orthogonal matrices, and $\Sigma \in \mathbb{R}^{n\times d}$ is a rectangular diagonal matrix. Substituting for S_1 in (D.39) from (D.40) and multiplying the result from left and right by U' and $V\Sigma'$, respectively, I get

$$\Sigma(\Sigma'\Sigma)^{-1}\Sigma' = \Sigma\Sigma' - X\Sigma\Sigma'X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma' - X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma'X\Sigma\Sigma' + \Sigma(\Sigma'\Sigma)^{-1}\Sigma'X\Sigma\Sigma'X'\Sigma(\Sigma'\Sigma)^{-1}\Sigma', \tag{D.41}$$

where $X \equiv U'\Gamma U$. Note that $\Sigma = \begin{pmatrix}\Sigma_1\\ \mathbf{0}\end{pmatrix}$ for some diagonal matrix $\Sigma_1 \in \mathbb{R}^{n\times d}$. Moreover, since $S_1'S_1$ is invertible, so is Σ_1. Therefore,

$$\Sigma(\Sigma'\Sigma)^{-1}\Sigma' = \begin{pmatrix} I & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix},$$

$$\Sigma\Sigma' = \begin{pmatrix} \Sigma_1^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}.$$

Write $X = \begin{pmatrix} X_{11} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}$, where $X_{11} \in \mathbb{R}^{d\times d}$, $X_{12} \in \mathbb{R}^{d\times(n-d)}$, $X_{21} \in \mathbb{R}^{(n-d)\times d}$, and $X_{22} \in \mathbb{R}^{(n-d)\times(n-d)}$. Equation (D.41) then implies

$$X_{11}'X_{11} = I - \Sigma_1^{-2}, \tag{D.42}$$

$$X_{21}\Sigma_1^2X_{11}' + X_{12}'X_{11}\Sigma_1^2 = 0. \tag{D.43}$$

These equations fully characterize the set of all (local) extrema of the KLDR.

I next use these equations to show that, as long as either d is equal to one or Γ_1 is symmetric, and for any $i = 1, \ldots, d$, the ith coordinate vector $e_i \in \mathbb{R}^n$ is an eigenvector of $(X + X')/2$ with eigenvalue $e_i'Xe_i$.[53] If $e_i'Xe_i = 0$, then trivially e_i is an eigenvector of $(X+X')/2$ with eigenvalue $e_i'Xe_i = 0$. So in the rest of the proof, I consider the case where $e_i'Xe_i \neq 0$. First, suppose $d = 1$.

[53] With slight abuse of notation, I use e_i to denote the ith coordinate vector both in $\mathbb{R}^n$ and in $\mathbb{R}^d$. Whether $e_i \in \mathbb{R}^n$ or $e_i \in \mathbb{R}^d$ will be clear from the context.

Then $i = 1$ and $X'_{11} = X_{11} = e'_1 X e_1 \neq 0$. On the other hand, Σ_1 is a non-zero scalar. Equation (D.43) then implies that $X_{21} + X'_{12} = 0$. Therefore,

$$\left(\frac{X+X'}{2}\right) e_1 = \frac{1}{2}\begin{pmatrix} 2X_{11} & X_{12}+X'_{21} \\ X_{21}+X'_{12} & X_{22}+X'_{22} \end{pmatrix}\begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} X_{11} \\ \mathbf{0} \end{pmatrix} = e'_1 X e_1 e_1,$$

proving that e_1 is an eigenvector of $(X+X')/2$ with eigenvalue $e_1 X e_1$. Next, suppose Γ_1 is symmetric. This implies that Γ, and by extension, X are symmetric matrices. Equation (D.42) then implies that X_{11} is a diagonal matrix. Since Σ_1 is also diagonal, it commutes with X_{11}. Equation (D.43) then implies that

$$(X_{21} + X'_{12})X_{11} = 2X_{21}X_{11} = \mathbf{0}, \tag{D.44}$$

where I am using the fact that Σ_1 is non-singular and X is symmetric. But since X_{11} is a diagonal matrix, it can be written as

$$X_{11} = \sum_{k=1}^{d} e'_k X_{11} e_k e_k e'_k.$$

Substituting in (D.44), I get

$$\sum_{k=1}^{d} X_{21} e'_k X_{11} e_k e_k e'_k = \mathbf{0}.$$

In particular, it must be the case that $X_{21} e'_i X_{11} e_i e_i = 0$. But since $e'_i X_{11} e_i = e'_i X e_i \neq 0$, it must be that $X_{21} e_i = 0$. Therefore,

$$\left(\frac{X+X'}{2}\right) e_i = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}\begin{pmatrix} e_i \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} X_{11} e_i \\ X_{21} e_i \end{pmatrix} = \begin{pmatrix} e'_i X_{11} e_i e_i \\ 0 \end{pmatrix} = e'_i X e_i e_i,$$

where the third equality relies on the fact that X_{11} is diagonal. This proves that e_i is an eigenvector of $(X+X')/2$ with eigenvalue $e'_i X e_i e_i$.

I next show that any matrices M and S_1 that satisfy the first-order optimality conditions (D.37) and (D.38) must be of the form

$$M = \sum_{i=1}^{d} a_i v_i v'_i, \tag{D.45}$$

$$S_1 = \sum_{i=1}^{d} \frac{1}{\sqrt{1-a_i^2}} u_i v'_i, \tag{D.46}$$

where $\{a_i\}_{i=1}^d$ are eigenvalues of C_1, $u_i \in \mathbb{R}^n$ denotes an eigenvector with eigenvalue a_i normalized such that $u'_i u_k = \mathbb{1}_{\{i=k\}}$ for all $i, k \in \{1, \ldots, d\}$, and $\{v_i\}_{i=1}^d$ is an orthonormal basis for $\mathbb{R}^d$. To see this, first note that equation (D.40) can be written as

$$S_1 = U\Sigma V' = U \sum_{i=1^d} \sigma_i e_i e'_i V',$$

where σ_i denotes the ith diagonal element of $\Sigma \in \mathbb{R}^{n\times d}$. I let $u_i \equiv U e_i$ and $v_i \equiv V e_i$. Since U and V are orthogonal matrices, $\{u_i\}_{i=1}^d$ is a set of orthonormal vectors and $\{v_i\}_{i=1}^d$ is an orthonormal basis for $\mathbb{R}^d$. Therefore, to show that S_1 takes the form given in (D.46), I only need to show that u_i is an eigenvector of C_1 with eigenvalue a_i and $\sigma_i = 1/\sqrt{1-a_i^2}$. Note that

$$C_1 u_i = \frac{1}{2}\left(\Gamma + \Gamma'\right) U e_i = \frac{1}{2} U U' \left(\Gamma + \Gamma'\right) U e_i = U \left(\frac{X + X'}{2}\right) e_i = U e_i' X e_i e_i = e_i' X e_i u_i,$$

where the fourth equality uses the fact that e_i is an eigenvector of $(X + X')/2$. Therefore, u_i is an eigenvector of C_1. On the other hand, multiplying equation (D.42) from left and right by e_i' and e_i, respectively, for $i \in \{1, \ldots, d\}$ and using the fact that X_{11} is diagonal, I get

$$\left(e_i' X_{11} e_i\right)^2 = 1 - \sigma_i^{-2}.$$

But

$$e_i' X_{11} e_i = e_i' X e_i = e_i' \left(\frac{X + X'}{2}\right) e_i = e_i' U \left(\frac{\Gamma + \Gamma'}{2}\right) U e_i = u_i' C_1 u_i = u_i' a_i u_i = a_i,$$

where a_i denotes the eigenvalue of C_1 with eigenvector u_i. Therefore, $\sigma_i = 1/\sqrt{1-a_i^2}$. Finally, recall that $M = S_1' \Gamma' S_1 (S_1' S_1)^{-1}$. By assumption, either $d = 1$, and so, S_1 is a vector in $\mathbb{R}^n$ or Γ is symmetric. Either way $S_1' \Gamma' S_1 = S_1'(\Gamma + \Gamma') S_1 / 2 = S_1' C_1 S_1$. Therefore,

$$\begin{aligned} M = S_1' C_1 S_1 (S_1' S_1)^{-1} &= \left(\sum_{i,k=1}^{d} \frac{1}{\sqrt{1-a_i^2}} v_i u_i' C_1 \frac{1}{\sqrt{1-a_k^2}} u_k v_k'\right) \left(\sum_{i,k=1}^{d} \frac{1}{\sqrt{1-a_i^2}} v_i u_i' \frac{1}{\sqrt{1-a_k^2}} u_k v_k'\right)^{-1} \\ &= \left(\sum_{i=1}^{d} \frac{1}{1-a_i^2} v_i a_i v_i'\right) \left(\sum_{i=1}^{d} \frac{1}{1-a_i^2} v_i v_i'\right)^{-1} = \sum_{i=1}^{d} a_i v_i v_i', \end{aligned}$$

where I am using the facts that u_i is an eigenvector of C_1 with eigenvalue a_i and that $\{u_i\}_{i=1}^d$ and $\{v_i\}_{i=1}^d$ are orthonormal sets of vectors.

Although any M and S_1 of the forms (D.45) and (D.46) satisfy the necessary optimality condition, not all such candidates are global minimizers of the KLDR. To find the global optima, I substitute the solutions to the first-order optimality conditions in the KLDR and select the solutions that minimize the KLDR. Multiplying equation (D.38) from left by S_1', I get

$$I = S_1' S_1 - S_1' \Gamma S_1 M - S_1' \Gamma' S_1 M' + S_1' S_1 M' M.$$

Computing the trace of the above equation and substituting the result in (D.36), I get

$$\begin{aligned} \text{KLDR}(M, S_1) &= -\frac{1}{2} \log\det\left(S_1' S_1\right) + \frac{1}{2} \operatorname{tr}\left(S_1' S_1\right) - \operatorname{tr}\left(M S_1' \Gamma S_1\right) + \frac{1}{2} \operatorname{tr}\left(M S_1' S_1 M'\right) + \text{constant} \\ &= -\frac{1}{2} \log\det\left(S_1' S_1\right) + \frac{1}{2} \operatorname{tr}(I) + \text{constant}. \end{aligned}$$

Therefore, the M and S_1 pairs that minimize the KLDR are the ones that maximize the determinant of $S_1' S_1$. But since $S_1' S_1$ is a symmetric matrix with eigenvalues $\{1/(1-a_i^2)\}_{i=1}^d$, its determinant is equal to $\prod_{i=1}^d \frac{1}{1-a_i^2}$. Therefore, any M and S_1 pair that minimize the KLDR are of the forms (D.45) and (D.46) with $\{a_i\}_{i=1}^d$ the top d eigenvalues of C_1 in magnitude (with the possibility that some of the a_i are equal).

With the expressions for the pseudo-true M and S_1 in hand, I can prove the theorem.

Part (a). The forecasts given a model parameterized by matrices M, D, and N are given by equation (D.17). Using the definition of $S \equiv \Gamma_0^{\frac{1}{2}} N$ and the fact that $D'D$ can be taken to be identity matrix, I can write equation (D.17) as follows:

$$E_t^{\theta}[y_{t+s}] = \Gamma_0^{\frac{1}{2}} S'^{-1} D M^s D' S' \Gamma_0^{\frac{-1}{2}} y_{t-\tau}.$$

Note that for any matrix $S = (S_1 \ S_2)$ that satisfies the first-order optimality condition with respect to S_2,

$$S^{-1} = \begin{pmatrix} (S_1'S_1)^{-1}S_1' \\ S_2' \end{pmatrix}.$$

Therefore,

$$S'^{-1} = \left(S_1(S_1'S_1)^{-1} \quad S_2\right),$$

and so

$$S'^{-1}D = S_1(S_1'S_1)^{-1}. \tag{D.47}$$

The forecasts can thus be written only in terms of matrices M and S_1 as follows:

$$E_t^{\theta}[y_{t+s}] = \Gamma_0^{\frac{1}{2}} S_1(S_1'S_1)^{-1} M^s S_1' \Gamma_0^{\frac{-1}{2}}.$$

Substituting for M and S_1 using (D.45) and (D.46) and simplifying the resulting expression, I get

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

Letting $p_i \equiv \Gamma_0^{\frac{-1}{2}} u_i$ and $q_i \equiv \Gamma_0^{\frac{1}{2}} u_i$ completes the proof of part (a).

Part (b). Equation (D.18) gives the variance-covariance matrix under a model parameterized by matrices M, D, and N. Using the definition of S and setting $D'D = I$, equation (D.18) can be written as follows:

$$\mathrm{Var}^{\theta}(y) = \Gamma_0^{\frac{1}{2}} \left(S'^{-1}S^{-1} + S^{-1\prime} D \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} D' S^{-1} \right) \Gamma_0^{\frac{1}{2}}.$$

To prove part (b), I need to show that the terms in parentheses add up to the identity matrix. I start with the first term:

$$S'^{-1}S^{-1} = (SS')^{-1} = (S_1S_1' + S_2S_2')^{-1}. \tag{D.48}$$

The fact that $S_2'S_2 = I$ implies that S_2 can be written as

$$S_2 = \sum_{i=d+1}^{n} u_i w_i',$$

where $u_i \in \mathbb{R}^n$ and $w_i \in \mathbb{R}^{n-d}$ for $i = d+1, \ldots, n$, $\{u_i\}_{i=d+1}^n$ are orthonormal vectors, and $\{w_i\}_{i=d+1}^n$ constitutes an orthonormal basis for $\mathbb{R}^{n-d}$. On the other hand, the fact that $S_1'S_2 = \mathbf{0}$ implies that

$u_i'u_k = 0$ for any $i \in \{1,\dots,d\}$ and $k \in \{d+1,\dots,n\}$. Therefore, $\{u_i\}_{i=1}^n$ constitutes an orthonormal basis for $\mathbb{R}^n$. Substituting for S_1 and S_2 in (D.48), I get

$$S'^{-1}S^{-1} = \left(\sum_{i=1}^{d} \frac{1}{1-a_i^2} u_i u_i' + \sum_{i=d+1}^{n} u_i u_i'\right)^{-1} = \sum_{i=1}^{d}(1-a_i^2)u_i u_i' + \sum_{i=d+1}^{n} u_i u_i',$$

where the second equality uses the fact that $\{u_i\}_{i=1}^n$ are orthonormal. Next consider the second term:

$$\begin{aligned}
S^{-1\prime} D \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} D' S^{-1} &= S_1 (S_1' S_1)^{-1} \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} (S_1' S_1)^{-1} S_1' \\
&= \sum_{i=1}^{d} \sqrt{1-a_i^2}\, u_i v_i' \sum_{\tau=1}^{\infty} \sum_{k=1}^{d} a_k^{2\tau} v_k v_k' \sum_{l=1}^{d} \sqrt{1-a_l^2}\, v_l u_l' \\
&= \sum_{i=1}^{d} (1-a_i^2) u_i u_i' \sum_{\tau=1}^{\infty} a_i^{2\tau} \\
&= \sum_{i=1}^{d} a_i^2 u_i u_i',
\end{aligned}$$

where the first equality uses (D.47) and the second equality is by (D.45) and (D.46). Putting everything together,

$$S'^{-1}S^{-1} + S^{-1\prime} D \sum_{\tau=1}^{\infty} M^{\tau} M'^{\tau} D' S^{-1} = \sum_{i=1}^{d}(1-a_i^2)u_i u_i' + \sum_{i=d+1}^{n} u_i u_i' + \sum_{i=1}^{d} a_i^2 u_i u_i' = \sum_{i=1}^{n} u_i u_i' = I,$$

where the last equality follows the fact that $\{u_i\}_{i=1}^n$ is an orthonormal basis for $\mathbb{R}^n$. □

Proof of Proposition 3

Recall that I have assumed (without loss of generality) that Γ_0 is non-singular. Since C_1 is symmetric, $\{u_i\}_{i=1}^d$ constitutes an orthonormal basis for $\mathbb{R}^n$, and so, $\Gamma_0^{\frac{-1}{2}} y_t$ 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 ith 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 ith largest eigenvalue of C_1 in magnitude. □

Proof of Proposition 4

I prove the result under the assumption that the top D eigenvalues of the first autocorrelation matrix, C_1, are all distinct. This assumption is true for generic true processes. By Theorem 5(a) (or Theorem 4), the forecasts of an agent who uses a pseudo-true d-state model θ are given by

$$E_t^\theta [y_{t+s}] = \sum_{i=1}^{d} a_i{}^s q_i p_i{}' y_t, \tag{D.49}$$

where a_i is the ith 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 multiplicative constants). Therefore, all agents use the same values of $\{(a_i, p_i, q_i)\}_i$ to forecast.

Consider agent j who is constrained to models of dimension d_j. The agent's optimal action given her pseudo-true d-state model is given by

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

where θ_j denotes agent j's pseudo-true model, $y_t^{(i)} \equiv p_i' y_t$ 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, $x_t \equiv (x_{1t}, \ldots, x_{Jt})' \in \mathbb{R}^J$, I can write the above expression as

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

where

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

$$g_j \equiv \begin{pmatrix} g_{j1} & g_{j2} & \dots g_{jd_j} & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{1\times D},$$

and

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

This completes the proof of the proposition. □

Proof of Proposition 5

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 a pseudo-true d-state model θ,

$$\begin{aligned}
E^\theta \left[v'y_t v'y_{t-l}\right] &= v'E^\theta\left[y_t y_{t-l}'\right] v = v'E^\theta\left[E_{t-l}^\theta\left[y_t\right] y_{t-l}'\right] v \\
&= v' \sum_{i=1}^{d} a_i{}^l q_i p_i' E^\theta\left[y_{t-l} y_{t-l}'\right] v = \sum_{i=1}^{d} a_i{}^l v' q_i p_i' \Gamma_0 v \\
&= \sum_{i=1}^{d} a_i{}^l v' \Gamma_0^{\frac{1}{2}} u_i u_i' \Gamma_0^{\frac{1}{2}} v = \sum_{i=1}^{d} a_i{}^l w' u_i u_i' w,
\end{aligned}$$

where the first equality is by Theorem 1, 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 5(a) (or Theorem 4 depending on the assumptions), and the fourth equality is by Theorem 5(b) (or Theorem 3). 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,

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

for any pseudo-true model θ. Furthermore,

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

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 5(b) (or Theorem 3), the variance of $y_t^{(1)}$ is the same under the true and pseudo-true 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,

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

for any pseudo-true model θ, 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 $y_t^{(n)}$. □

Proof of Proposition 6

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 4,

$$E_t^{\theta^*} \left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta} \hat{x}_{t+s} - \sigma \left(\hat{i}_{t+s} - r_{t+s}^n \right) - \frac{\sigma}{\beta} \hat{\pi}_{t+s} \right) \right] = \gamma_x \hat{z}_t, \tag{D.50}$$

$$E_t^{\theta^*} \left[\sum_{s=1}^{\infty} (\beta\delta)^s \left(\kappa \hat{x}_{t+s} + \frac{1-\delta}{\delta} \hat{\pi}_{t+s} + \mu_{t+s} \right) \right] = \gamma_\pi \hat{z}_t, \tag{D.51}$$

where γ_x and γ_π are constants that are to be determined in equilibrium, $\hat{z}_t = p' f_t$ is agents' time-t estimate of the subjective state, and $p \equiv (p_x, p_\pi, p_i)'$ is the relative attention vector. Substituting in (10) and (11) and collecting terms, I get

$$\sigma r_t^n = \hat{x}_t + \sigma \hat{i}_t - \gamma_x \left(p_x \hat{x}_t + p_\pi \hat{\pi}_t + p_i \hat{i}_t \right), \tag{D.52}$$

$$\mu_t = \hat{\pi}_t - \kappa \hat{x}_t - \gamma_\pi \left(p_x \hat{x}_t + p_\pi \hat{\pi}_t + p_i \hat{i}_t \right). \tag{D.53}$$

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 (D.50) and (D.51) and using Theorem 4 to characterize the resulting subjective expectations, I get

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

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

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 (D.52) and (D.53) for $\hat{x}_t$ and $\hat{\pi}_t$ results in equations (12) and (13). □

Proof of Proposition 7

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 4,

$$E_t^{\theta^*}\left[\sum_{s=1}^{\infty}\beta^s\left(\frac{1-\beta}{\beta}\hat{x}_{t+s} - \sigma\left(\hat{i}_{t+s} - r_{t+s}^n\right) - \frac{\sigma}{\beta}\hat{\pi}_{t+s}\right)\right] = \iota_x r_t^n, \tag{D.54}$$

$$E_t^{\theta^*}\left[\sum_{s=1}^{\infty}(\beta\delta)^s\left(\kappa\hat{x}_{t+s} + \frac{1-\delta}{\delta}\hat{\pi}_{t+s} + \mu_{t+s}\right)\right] = \iota_\pi r_t^n, \tag{D.55}$$

for some constants ι_x and ι_π, which are to be determined in equilibrium. Substituting in (10) and (11) and collecting terms, I get

$$\hat{x}_t = \iota_x r_t^n, \tag{D.56}$$

$$\hat{\pi}_t = \kappa\hat{x}_t + \iota_\pi r_t^n = (\kappa\iota_x + \iota_\pi)r_t^n. \tag{D.57}$$

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 (D.54) and (D.55) and using Theorem 4 to characterize the resulting subjective expectations, I get

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

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

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. □

Proof of Proposition B.1

I first prove a useful lemma, which offers a canonical representation of the autocorrelation matrices for stochastic processes that can be represented as in (B.1):[54]

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

Lemma D.5. *Suppose $\{C_l\}_l$ are the autocorrelation matrices of a non-degenerate n-dimensional stationary ergodic process that can be represented as in (B.1) 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*

$$C_l = \mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H}. \tag{D.58}$$

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 (D.58), which can be represented as in (B.1).[55]

Proof. The assumption that the process is non-degenerate requires $m \geq n$, an assumption I maintain throughout the first part of the proof. Given representation (B.1), the autocovariance matrices are given by

$$\Gamma_l = \mathbb{E}\left[y_t y'_{t-l}\right] = H' F^l \mathbb{E}\left[f_{t-l} f'_{t-l}\right] H = H' F^l V H,$$

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

$$V = F V F' + \Sigma, \tag{D.59}$$

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

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

Matrix V is positive semidefinite; it is positive definite if the representation in (B.1) is minimal.[56] Without loss of generality, I assume that that is the case. Define

$$\mathbb{H}' \equiv (H'VH)^{\frac{-1}{2}} H' V^{\frac{1}{2}},$$
$$\mathbb{F} \equiv V^{\frac{-1}{2}} F V^{\frac{1}{2}}.$$

Then

$$C_l = \mathbb{H}' \left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2} \right) \mathbb{H}. \tag{D.60}$$

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 (D.59), I get

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

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,

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

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

[55] Matrix $\mathbb{H} \in \mathbb{R}^{m \times n}$ is semi-orthogonal if $\mathbb{H}'\mathbb{H} = I$, where I denotes the $n \times n$ identity matrix.
[56] See, for instance, Akaike (1975).

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 (D.60) 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 (D.59) 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 (B.1) with matrices F, H, and Σ as above are given by (D.60) with $\mathbb{F} = \hat{\mathbb{F}}$ and $\mathbb{H} = \hat{\mathbb{H}}$. □

Proof of Proposition B.1. By Lemma D.5,

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

where $\mathbb{H}' \equiv (H'VH)^{\frac{-1}{2}} H'V^{\frac{1}{2}}$, $\mathbb{F} \equiv V^{\frac{-1}{2}} F V^{\frac{1}{2}}$, and $V \equiv \mathbb{E}\left[f_t f_t'\right]$ 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 D.5, $\mathbb{H} = H$ is an orthogonal matrix. Thus,

$$\rho(C_l) = \rho\left(\mathbb{H}'\left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2}\right)\mathbb{H}\right) = \rho\left(\frac{\mathbb{F}^l + \mathbb{F}'^l}{2}\right) = \rho\left(\frac{F^l + F'^l}{2}\right) \tag{D.61}$$

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

$$\rho\left(\frac{F^l + F'^l}{2}\right) = \left\|\frac{F^l + F'^l}{2}\right\|_2 \leq \frac{1}{2}\left\|F^l\right\|_2 + \frac{1}{2}\left\|F'^l\right\|_2 = \left\|F^l\right\|_2 \leq \|F\|_2^l. \tag{D.62}$$

Therefore,

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

On the other hand, by equations (D.61) and (D.62),

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

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. □

Proof of Proposition B.2

I first state and prove a useful lemma:

Lemma D.6. *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 $u'u = 1$.*[57] *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 Lemma D.3. As in the proof of Lemma D.3, I present the argument under the assumption that the largest eigenvalue of $C(a, \eta)$ is simple at the point (a^*, η^*) that maximizes $\lambda_{\max}(C(a, \eta))$.[58] I start by proposing a candidate solution to the problem 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 (D.30) and (D.31), I get

$$\left.\frac{\partial \lambda_{\max}(\Omega(a, \eta))}{\partial a}\right|_{\eta=0} = -2a + 2u'_{\max}(aC_1)C_1 u_{\max}(aC_1),$$
$$\left.\frac{\partial \lambda_{\max}(\Omega(a, \eta)))}{\partial \eta}\right|_{\eta=0} = 2a^2 - 2(1 + a^2)\lambda_{\max}(aC_1) + 2a^2 u'_{\max}(aC_1)C_2 u_{\max}(aC_1),$$

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

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

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. □

Proof of Proposition B.2. Let σ^2 denote the variance of y_t. By the argument in the proof of Lemma D.5, the lag-l autocorrelation of y_t is given by

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

where $\mathbb{F} \equiv V^{\frac{-1}{2}} F V^{\frac{1}{2}}$, $\mathbb{H}' \equiv (H'VH)^{\frac{-1}{2}} H'V^{\frac{1}{2}}$, and V is the solution to the discrete-time Lyapunov equation (D.59). Since F and Σ are diagonal matrices, so is V. Therefore, $\mathbb{F} = F$. On the other hand,

[57] The 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 $|\lambda| = \rho(C_1)$.

[58] See footnote 50 for how the argument can be generalized.

by Lemma D.5, $\mathbb{H}$ is a semi-orthogonal matrix. Therefore, $\mathbb{H}'\mathbb{H} = 1$, and so,

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

where $w_i \equiv \mathbb{H}_i^2 \geq 0$, $\sum_{i=1}^m w_i = 1$, and α_i is the ith diagonal element of 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 (B.1) 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 $\widetilde{\mathbb{P}}$ represented as in (B.1) with $F = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_{\tilde{m}})$, $\epsilon_t \sim \mathcal{N}(0, \Sigma)$, $\Sigma = I - FF'$, and $H = \sigma\,\text{diag}(\sqrt{\tilde{w}_1}, \dots, \sqrt{\tilde{w}_{\tilde{m}}})$. By the argument in the proof of Lemma D.5, $\widetilde{\mathbb{P}}$ has the same autocorrelation matrices as $\mathbb{P}$. Moreover, both $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are mean-zero and normal and both have variance σ^2. Therefore, $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ 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, $u'C_2u = C_2 > C_1^2 = \rho(C_1)^2$, where I am using the fact that y_t is a scalar. Thus, by Lemma D.6, $\eta > 0$. To see why $\eta < 1$, recall that by Theorem 2, 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 Lemma D.3. The proposition then follows Theorem 2 by noting that $qp' = 1$ whenever y_t is a scalar. □

E Omitted Details for the NK Application

E.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 (D.52) and (D.53) in (10) and (11), I get

$$E_t^{\theta^*}\left[\sum_{s=1}^{\infty} \beta^s \left(\frac{1-\beta}{\beta}\hat{x}_{t+s} - \sigma\left(\hat{i}_{t+s} - r_{t+s}^n\right) - \frac{\sigma}{\beta}\hat{\pi}_{t+s}\right)\right] = E_t^{\theta^*}\left[\sum_{s=1}^{\infty} \beta^s v_x' f_{t+s}\right], \tag{E.1}$$

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

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

$$v_x' f_t = \frac{1}{\beta}\left[(1-\beta\gamma_x p_x)\hat{x}_t - (\sigma + \beta\gamma_x p_\pi)\hat{\pi}_t - \beta\gamma_x p_i \hat{i}_t\right],$$

$$v_\pi' f_t = \frac{1}{\delta}\left[-\delta\gamma_\pi p_x \hat{x}_t + (1-\delta\gamma_\pi p_\pi)\hat{\pi}_t - \delta\gamma_\pi p_i \hat{i}_t\right].$$

On the other hand,

$$E_t^{\theta^*}[f_{t+s}] = \Sigma_{f_s\omega_T}\Sigma_{\omega_T\omega_T}^{-1}\omega_T, \tag{E.3}$$

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

$$E_t^{\theta^*}\left[\sum_{s=1}^{\infty}\beta^s v_x' f_{t+s}\right] = \psi_{xT}'\omega_T,$$
$$E_t^{\theta^*}\left[\sum_{s=1}^{\infty}(\beta\delta)^s v_\pi' f_{t+s}\right] = \psi_{\pi T}'\omega_T.$$

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

$$\psi_{xT}' \equiv (\psi_{xf}', \psi_{xi_1}, \ldots, \psi_{xi_T})' \equiv v_x'\left(\sum_{s=1}^{\infty}\beta^s\Sigma_{f_s\omega_T}\right)\Sigma_{\omega_T\omega_T}^{-1}, \tag{E.4}$$

$$\psi_{\pi T}' \equiv (\psi_{\pi f}', \psi_{\pi i_1}, \ldots, \psi_{\pi i_T})' \equiv v_\pi'\left(\sum_{s=1}^{\infty}(\beta\delta)^s\Sigma_{f_s\omega_T}\right)\Sigma_{\omega_T\omega_T}^{-1}, \tag{E.5}$$

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,

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

Substituting in equations (10) and (11), I get

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

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

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

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 4,

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

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

$$E^{\theta^*}[f_{t+s}\hat{i}_{t+\tau}] = E^{\theta^*}\left[f_{t+s}E^{\theta^*}[\hat{i}_{t+\tau}|f_{t+s}]\right] = E^{\theta^*}[f_{t+s}a^{\tau-s}q_i p' f_{t+s}] = a^{\tau-s}q_i E^{\theta^*}\left[f_{t+s}f'_{t+s}\right]p = a^{\tau-s}q_i\Gamma_0 p.$$

Likewise, if $s>\tau$, then

$$E^{\theta^*}[f_{t+s}\hat{i}_{t+\tau}] = E^{\theta^*}\left[\hat{i}_{t+\tau}E^{\theta^*}[f_{t+s}|f_{t+\tau}]\right] = E^{\theta^*}\left[e'_i f_{t+\tau}a^{s-\tau}qp' f_{t+\tau}\right] = a^{s-\tau}qp' E^{\theta^*}\left[f_{t+s}f'_{t+s}\right]e_i = a^{s-\tau}qp'\Gamma_0 e_i,$$

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

$$E^{\theta^*}[f_{t+s}\hat{i}_{t+\tau}] = E^{\theta^*}[f_{t+s}f'_{t+s}e_i] = \Gamma_0 e_i.$$

I next compute the elements of $\Sigma_{\omega_T\omega_T}$. First, note that

$$E^{\theta^*}[f_t f'_t] = \Gamma_0,$$

and

$$E^{\theta^*}[f'_t\hat{i}_{t+\tau}] = E^{\theta^*}\left[f'_t E^{\theta^*}[\hat{i}_{t+\tau}|f_t]\right] = E^{\theta^*}\left[a^\tau q_i p' f_t f'_t\right] = a^\tau q_i p'\Gamma_0.$$

Finally, if $\tau<\tau'$, then

$$E^{\theta^*}[\hat{i}_{t+\tau}\hat{i}_{t+\tau'}] = E^{\theta^*}\left[\hat{i}_{t+\tau}E^{\theta^*}[\hat{i}_{t+\tau'}|f_{t+\tau}]\right] = E^{\theta^*}\left[e'_i f_{t+\tau}a^{\tau'-\tau}q_i p' f_{t+\tau}\right] = a^{\tau'-\tau}q_i p'\Gamma_0 e_i,$$

and

$$E^{\theta^*}[\hat{i}_{t+\tau}\hat{i}_{t+\tau}] = e'_i E^{\theta^*}[f_{t+\tau}f'_{t+\tau}]e_i = e'_i\Gamma_0 e_i.$$

Putting everything together, I get

$$\Sigma_{\omega_T\omega_T} = \begin{pmatrix} \Gamma_0 & aq_i\Gamma_0 p & a^2q_i\Gamma_0 p & \dots & a^Tq_i\Gamma_0 p \\ aq_ip'\Gamma_0 & e'_i\Gamma_0 e_i & aq_ip'\Gamma_0 e_i & \dots & a^{T-1}q_ip'\Gamma_0 e_i \\ a^2q_ip'\Gamma_0 & aq_ip'\Gamma_0 e_i & e'_i\Gamma_0 e_i & \dots & a^{T-2}q_ip'\Gamma_0 e_i \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^Tq_ip'\Gamma_0 & a^{T-1}q_ip'\Gamma_0 e_i & a^{T-2}q_ip'\Gamma_0 e_i & \dots & e'_i\Gamma_0 e_i \end{pmatrix}. \tag{E.6}$$

and

$$\Sigma_{f_s\omega_T} = \begin{cases} \begin{pmatrix} a^sqp'\Gamma_0 & \Gamma_0 e_i & aq_i\Gamma_0 p & a^2q_i\Gamma_0 p & \dots & a^{T-1}q_i\Gamma_0 p\end{pmatrix} & \text{if } s=1, \\ \begin{pmatrix} a^sqp'\Gamma_0 & a^{s-1}qp'\Gamma_0 e_i & \dots & aqp'\Gamma_0 e_i & \Gamma_0 e_i & aq_i\Gamma_0 p & \dots & a^{T-s}q_i\Gamma_0 p\end{pmatrix} & \text{if } 1<s<T, \\ \begin{pmatrix} a^sqp'\Gamma_0 & a^{s-1}qp'\Gamma_0 e_i & \dots & aqp'\Gamma_0 e_i & \Gamma_0 e_i\end{pmatrix} & \text{if } s=T, \\ \begin{pmatrix} a^sqp'\Gamma_0 & a^{s-1}qp'\Gamma_0 e_i & \dots & a^{s-T}qp'\Gamma_0 e_i\end{pmatrix} & \text{if } s>T. \end{cases} \tag{E.7}$$

Therefore,

$$\sum_{s=1}^{\infty}\beta^s\Sigma_{f_s\omega_T}$$
$$= \left(\sum_{s=1}^{\infty}(a\beta)^sqp'\Gamma_0 \quad \beta\Gamma_0 e_i + \sum_{s=2}^{\infty}a^{s-1}\beta^sqp'\Gamma_0 e_i \quad \dots \quad \sum_{s=1}^{T-1}a^{T-s}\beta^sq_i\Gamma_0 p + \beta^T\Gamma_0 e_i + \sum_{s=T+1}^{\infty}a^{s-T}\beta^sqp'\Gamma_0 e_i\right)$$

$$= \left(\frac{a\beta q p' \Gamma_0}{1-a\beta} \quad \beta\Gamma_0 e_i + \frac{a\beta^2 q p' \Gamma_0 e_i}{1-a\beta} \quad a\beta q_i \Gamma_0 p + \beta^2 \Gamma_0 e_i + \frac{a\beta^3 q p' \Gamma_0 e_i}{1-a\beta} \quad \dots \quad \frac{(a^T\beta - \beta^T a) q_i \Gamma_0 p}{a-\beta} + \beta^T \Gamma_0 e_i + \frac{a\beta^{T+1} q p' \Gamma_0 e_i}{1-a\beta} \right).$$

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 (E.4) and (E.5) to find ψ_{xT} and $\psi_{\pi T}$.

E.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 E.1 plots $\rho(C_l)$ and $\rho(C_1)^l$, where $\rho(C_l)$ denotes the spectral radius of the lag-l autocorrelation matrix of f_t.[59] The fact that the $\rho(C_l)$ line lies below the $\rho(C_1)^l$ line indicates that the estimated process is exponentially ergodic.

F Omitted Details for the RBC Application

F.1 Temporary Equilibrium

The (log-)linearized temporary-equilibrium conditions are given by

$$\hat{o}_t = \hat{a}_t + \alpha\hat{k}_t + (1-\alpha)\hat{n}_t, \tag{F.1}$$

$$\hat{w}_t = \hat{a}_t + \alpha(\hat{k}_t - \hat{n}_t), \tag{F.2}$$

$$\hat{r}_t = r\hat{a}_t + (1-\alpha) r (\hat{n}_t - \hat{k}_t), \tag{F.3}$$

$$\hat{n}_t = \frac{1}{\varphi}\hat{w}_t - \frac{1}{\sigma\varphi}\hat{c}_t, \tag{F.4}$$

$$\hat{i}_t = \frac{o}{i}\hat{y}_t - \frac{c}{i}\hat{c}_t, \tag{F.5}$$

$$\hat{k}_t = (1-\delta)\hat{k}_{t-1} + \delta\hat{i}_{t-1}, \tag{F.6}$$

$$\hat{a}_t = \rho\hat{a}_{t-1} + \epsilon_t, \tag{F.7}$$

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - \sigma\beta E_t[\hat{r}_{t+1}], \tag{F.8}$$

[59]The result would be identical if I instead used the autocorrelation matrices of s_t. This is a corollary of the linear-invariance result.

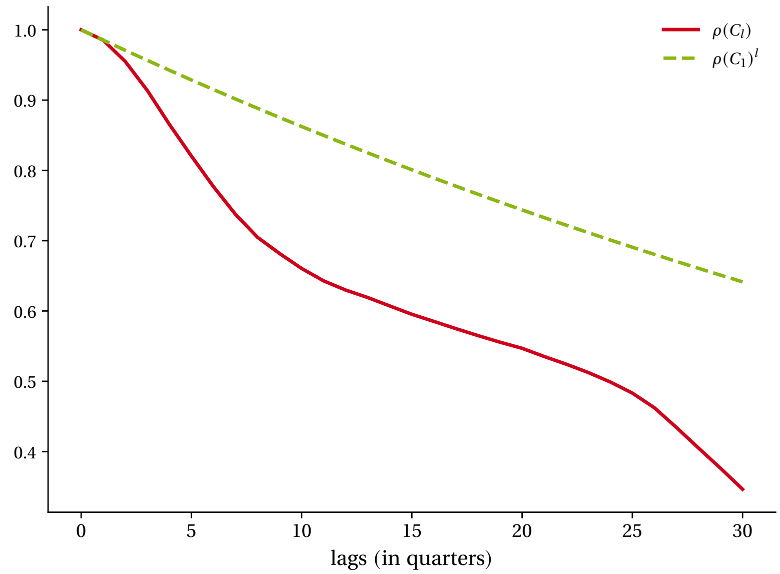

Figure E.1. Exponential ergodicity of the estimated new-Keynesian model

Notes: This figure plots $\rho(C_l)$ and $\rho(C_1)^l$ as functions of lag l, where $\rho(C_l)$ denotes the estimated spectral radius of the lag-l autocorrelation matrix of vector $f_t = (\hat{x}_t, \hat{\pi}_t, \hat{\imath}_t)'$, with $\hat{x}_t$ the percentage difference between Real GDP and Potential Output in period t, $\hat{\pi}_t$ the percentage change in GDP Deflator, and $\hat{\imath}_t$ the Effective Fed Funds Rate. Quarterly (unfiltered) U.S. data from the first quarter of 1955 to the fourth quarter of 2008.

where $\hat{r}_t$ denotes the first-order deviation of the interest rate from its steady-state value and the remaining hatted variables are log-deviations from the corresponding steady-state values. The Euler equation (F.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 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

$$\frac{1-\beta}{\chi}\sum_{s=0}^{\infty}\beta^s E_t[\hat{c}_{t+s}] = \frac{1}{\beta}\hat{k}_t + \sum_{s=0}^{\infty}\beta^s E_t[\hat{r}_{t+s}] + \zeta\sum_{s=0}^{\infty}\beta^s E_t[\hat{w}_{t+s}]. \tag{F.9}$$

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 (F.9), I get

$$\hat{c}_t = \frac{\chi}{\beta}\hat{k}_t + \chi\hat{r}_t + \chi\zeta\hat{w}_t + (\chi-\beta\sigma)\sum_{s=1}^{\infty}\beta^s E_t[\hat{r}_{t+s}] + \chi\zeta\sum_{s=1}^{\infty}\beta^s E_t[\hat{w}_{t+s}]. \tag{F.10}$$

F.2 Constrained-Rational-Expectations Equilibrium

Suppose households use a pseudo-true one-state model to forecast the wage rate and interest rate. Define $\omega_t \equiv (\hat{o}_t, \hat{n}_t, \hat{w}_t, \hat{r}_t, \hat{c}_t, \hat{i}_t)$, $f_t \equiv (\hat{k}_t, \hat{a}_t)'$, 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 (F.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,

$$\hat{c}_t = \frac{\chi}{\beta}\hat{k}_t + \chi\hat{r}_t + \chi\zeta\hat{w}_t + \sum_{s=1}^{\infty}\beta^s v' T E_t[f_{t+s}].$$

Households' forecasts of f_t when they use model θ is given by (4). This can be written recursively as

$$E_t^{\theta}[f_{t+s}] = a^s(1-\eta)\hat{z}_t q, \tag{F.11}$$

$$\hat{z}_t = a\eta\hat{z}_{t-1} + p'f_t = a\eta\hat{z}_{t-1} + p_k\hat{k}_t + p_a\hat{a}_t, \tag{F.12}$$

where $\hat{z}_t$ denote households' estimate of the subjective state at time t. Therefore,

$$\hat{c}_t = \frac{\chi}{\beta}\hat{k}_t + \chi\hat{r}_t + \chi\zeta\hat{w}_t + \frac{a\beta(1-\eta)}{1-a\beta}v'Tq\hat{z}_t. \tag{F.13}$$

I guess that $\eta = 0$ in equilibrium and later verify this guess. Solving for $\hat{z}_t$ from (F.12) and substituting in (F.13), I get

$$\hat{c}_t = \left(\frac{\chi}{\beta} + \gamma_k\right)\hat{k}_t + \chi\hat{r}_t + \chi\zeta\hat{w}_t + \gamma_a\hat{a}_t, \tag{F.14}$$

where

$$\gamma_k \equiv \frac{a\beta}{1-a\beta}v'Tqp_k, \tag{F.15}$$

$$\gamma_a \equiv \frac{a\beta}{1-a\beta}v'Tqp_a. \tag{F.16}$$

Equations (F.1)–(F.5) and (F.14) can be solved for ω_t as a function of f_t. This verifies the guess that $\xi_t = (f_t', \omega_t')' = Tf_t$ and leads to an expression for matrix T. In particular,

$$\hat{i}_t = \psi_k\hat{k}_t + \psi_a\hat{a}_t,$$

for some ψ_k and ψ_a. Substituting for $\hat{i}_{t-1}$ from above in (F.6), I get

$$\hat{k}_t = (1-\delta+\delta\psi_k)\hat{k}_{t-1} + \delta\psi_a\hat{a}_{t-1}. \tag{F.17}$$

I can now describe the constrained-rational-expectations equilibrium. Equations (F.7) and (F.17) can be written in vector form as

$$f_t = \mathbb{F}(\gamma_k, \gamma_a)f_{t-1} + \epsilon_t. \tag{F.18}$$

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 (F.18) with $\gamma_k = \gamma_k^*$ and $\gamma_a = \gamma_a^*$, (ii) γ_k^* and γ_a^* are given by equations (F.15) and (F.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 (F.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 (F.15) and (F.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.

G Omitted Details for the DMP Application

G.1 Timing

Each period is divided into three sub-periods:

1. Potential employers simultaneously decide whether to post a vacancy.
2. Workers and vacancies match, and separations happen.
3. Production takes place, wages are paid, and workers consume.

A value function indexed with subscript t represents the value at the beginning of period t.

G.2 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:

$$U_t = p_t \left(w_t + \beta E_t [V_{t+1}]\right) + (1 - p_t) \left(b + \beta E_t [U_{t+1}]\right), \tag{G.1}$$

$$V_t = s_t \left(b + \beta E_t [U_{t+1}]\right) + (1 - s_t) \left(w_t + \beta E_t \left[V_{t+1}\right]\right), \tag{G.2}$$

where b denotes 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

$$V_t - U_t = (1 - s_t - p_t) \left(w_t - b + \beta E_t \left[V_{t+1} - U_{t+1}\right]\right). \tag{G.3}$$

Define

$$\lambda^w_{t,t+\tau} \equiv \prod_{k=0}^{\tau} (1 - s_{t+k} - p_{t+k}). \tag{G.4}$$

Solving (G.3) forward, I get

$$V_t - U_t = \lambda^w_{t,t}(w_t - b) + E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau \lambda^w_{t,t+\tau}(w_{t+\tau} - b)\right]. \tag{G.5}$$

This equation is valid under arbitrary expectations.

I consider the firms' problem next. Let J_t denote the time-t value to a firm of a job. It solves the following Bellman equation:

$$J_t = (1 - s_t) \left(a_t - w_t + \beta E_t [J_{t+1}]\right).$$

Solving the equation forward, I get

$$J_t = \lambda^f_{t,t}(a_t - w_t) + E_t \left[\sum_{\tau=1}^{\infty} \beta^\tau \lambda^f_{t,t+\tau}(a_{t+\tau} - w_{t+\tau})\right], \tag{G.6}$$

where

$$\lambda^f_{t,t+\tau} \equiv \prod_{k=0}^{\tau} (1 - s_{t+k}). \tag{G.7}$$

Free entry by firms implies

$$0 = -k + q_t \left(a_t - w_t + \beta E_t [J_{t+1}]\right), \tag{G.8}$$

where $q_t = \mu\theta_t^{-\alpha}$ is the probability of filling a vacancy in each period. Substituting for J_t in (G.8) from (G.6), I get

$$\theta_t^\alpha = \frac{\mu}{k}(a_t - w_t) + \frac{\mu}{k}E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\lambda^f_{t+1,t+\tau}(a_{t+\tau} - w_{t+\tau})\right]. \tag{G.9}$$

Equation (G.9) determines tightness as a function of the current labor productivity and the wage as well as firms' expectations of those variables.

The wage rate is determined by Nash bargaining. Under Nash bargaining,

$$\frac{a_t - w_t + \beta E_t[J_{t+1}]}{1-\delta} = \frac{w_t - b + \beta E[V_{t+1} - U_{t+1}]}{\delta},$$

where δ denotes workers' bargaining power. Combining the above equation with (G.5) and (G.6) and solving for w_t, I get

$$w_t = \delta a_t + (1-\delta)b + \delta E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\lambda^f_{t+1,t+\tau}(a_{t+\tau} - w_{t+\tau})\right] - (1-\delta)E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau\lambda^w_{t+1,t+\tau}(w_{t+\tau} - b)\right]. \tag{G.10}$$

The unemployment rate follows the first-order difference equation

$$u_t = u_{t-1} + s_{t-1}(1 - u_{t-1}) - p_{t-1}u_{t-1}. \tag{G.11}$$

G.3 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 (G.10) 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 (G.6) implies that the value of a job to a firm is constant in the steady state:

$$J = \frac{1-s}{1-\beta(1-s)}(1-w).$$

Equation (G.8) and the definition of q_t imply

$$\frac{\mu}{k\theta^\alpha} = \frac{1-\beta(1-s)}{1-w} = \frac{1-s}{J}.$$

The steady-state unemployment rate satisfies

$$s\frac{1-u}{u} = p.$$

G.4 Log-Linear Model

I next log-linearize the model around the steady state. Log-linearizing (G.4) and (G.7), I get

$$\hat{\lambda}^w_{t,t+\tau} = -\frac{1}{1-s-p}\sum_{k=0}^{\tau}\left(p\hat{p}_{t+k} + s\hat{s}_{t+k}\right),$$
$$\hat{\lambda}^f_{t,t+\tau} = -\frac{1}{1-s}\sum_{k=0}^{\tau} s\hat{s}_{t+k}.$$

Log-linearizing $p_t = \mu\theta_t^{1-\alpha}$, I get

$$\hat{p}_t = (1-\alpha)\hat{\theta}_t.$$

Log-linearizing (G.9),

$$\hat{\theta}_t = \frac{\mu}{\alpha k\theta^\alpha}\left((1-b)\hat{a}_t - w\hat{w}_t + E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau\left((1-b)\hat{a}_{t+\tau} - w\hat{w}_{t+\tau} + (1-w)\hat{\lambda}^f_{t+1,t+\tau}\right)\right]\right).$$

The term involving $\hat{\lambda}^f_{t,t+\tau}$ can be simplified further:

$$\begin{aligned}
\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau\hat{\lambda}^f_{t+1,t+\tau} &= -\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau\frac{s}{1-s}\sum_{k=0}^{\tau-1}\hat{s}_{t+1+k} \\
&= -s\sum_{k=0}^{\infty}\hat{s}_{t+1+k}\sum_{\tau=k+1}^{\infty}\beta^\tau(1-s)^{\tau-1} \\
&= -\frac{\beta s}{1-\beta(1-s)}\sum_{\tau=1}^{\infty}\beta^{\tau-1}(1-s)^{\tau-1}\hat{s}_{t+\tau}.
\end{aligned}$$

Define

$$\zeta \equiv \frac{s(1-w)}{(1-s)(1-\beta(1-s))}.$$

Then,

$$\hat{\theta}_t = \frac{1-s}{\alpha J}\left((1-b)\hat{a}_t - w\hat{w}_t + E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau\left((1-b)\hat{a}_{t+\tau} - w\hat{w}_{t+\tau} - \zeta\hat{s}_{t+\tau}\right)\right]\right), \tag{G.12}$$

where I am using the fact that $\frac{\mu}{k\theta^\alpha} = \frac{1-s}{J}$. Log-linearizing (G.10),

$$\begin{aligned}
w\hat{w}_t = \delta(1-b)\hat{a}_t + \delta E_t&\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau((1-b)\hat{a}_{t+\tau} - w\hat{w}_{t+\tau} + (1-w)\hat{\lambda}^f_{t+1,t+\tau})\right] \\
-(1-\delta)E_t&\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau\left((w-b)\hat{\lambda}^w_{t+1,t+\tau} + w\hat{w}_{t+\tau}\right)\right].
\end{aligned} \tag{G.13}$$

The terms involving $\hat{\lambda}^w_{t+1,t+\tau}$ and $\hat{\lambda}^f_{t+1,t+\tau}$ can be simplified further:

$$\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau\hat{\lambda}^w_{t,t+\tau} = -\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau\frac{1}{1-s-p}\sum_{k=0}^{\tau-1}\left(p\hat{p}_{t+1+k} + s\hat{s}_{t+1+k}\right)$$

$$= -\beta \sum_{k=0}^{\infty} (p\hat{p}_{t+1+k} + s\hat{s}_{t+1+k}) \sum_{\tau=k+1}^{\infty} (\beta(1-s-p))^{\tau-1}$$

$$= -\frac{\beta}{1-\beta(1-s-p)} \sum_{\tau=1}^{\infty} \beta^{\tau-1}(1-s-p)^{\tau-1} (p\hat{p}_{t+\tau} + s\hat{s}_{t+\tau}) .$$

$$\sum_{\tau=1}^{\infty} \beta^{\tau}(1-s)^{\tau} \hat{\lambda}^{f}_{t+1,t+\tau} = -\frac{\beta s}{1-\beta(1-s)} \sum_{\tau=1}^{\infty} \beta^{\tau-1}(1-s)^{\tau-1} \hat{s}_{t+\tau}.$$

Define

$$\chi \equiv \frac{(1-\delta)(w-b)}{(1-s-p)(1-\beta(1-s-p))}$$

Then, (G.13) can be written as

$$w\hat{w}_t = \delta(1-b)\hat{a}_t + E_t\left[\sum_{\tau=1}^{\infty} \beta^{\tau}(1-s)^{\tau}(\delta(1-b)\hat{a}_{t+\tau} - \delta w\hat{w}_{t+\tau} - \delta\zeta\hat{s}_{t+\tau})\right]$$
$$- E_t\left[\sum_{\tau=1}^{\infty} \beta^{\tau}(1-s-p)^{\tau}\left((1-\delta)w\hat{w}_{t+\tau} - p\chi(1-\alpha)\hat{\theta}_{t+\tau} - s\chi\hat{s}_{t+\tau}\right)\right]. \tag{G.14}$$

Finally, log-linearizing (G.11),

$$\hat{u}_t = (1-s-p)\hat{u}_{t-1} - (1-\alpha)p\hat{\theta}_{t-1} + p\hat{s}_{t-1}. \tag{G.15}$$

G.5 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 (G.12) and (G.14), I get

$$\hat{\theta}_t = \frac{1-s}{1-\beta\rho_a(1-s)}\frac{1-b-\gamma_{wa}}{\alpha J}\hat{a}_t - \frac{1-s}{1-\beta\rho_s(1-s)}\frac{\beta\rho_s(1-s)\zeta+\gamma_{ws}}{\alpha J}\hat{s}_t,$$

and

$$w\hat{w}_t = \left[\delta(1-b) + \frac{\beta\delta\rho_a(1-s)(1-b-\gamma_{wa})}{1-\beta\rho_a(1-s)} + \frac{\beta\rho_a(1-s-p)}{1-\beta\rho_a(1-s-p)}(p\chi(1-\alpha)\gamma_{\theta a} - (1-\delta)\gamma_{wa})\right]\hat{a}_t$$
$$+ \left[\frac{\beta\rho_s(1-s-p)}{1-\beta\rho_s(1-s-p)}(p\chi(1-\alpha)\gamma_{\theta s} + s\chi - (1-\delta)\gamma_{ws}) - \frac{\beta\delta\rho_s(1-s)(\zeta+\gamma_{ws})}{1-\beta\rho_s(1-s)}\right]\hat{s}_t.$$

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 (17) and (G.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.

G.6 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,

$$\hat{\theta}_t = \psi_{\theta u}\hat{u}_t + \psi_{\theta a}\hat{a}_t + \psi_{\theta s}\hat{s}_t,$$

$$w\hat{w}_t = \psi_{wu}\hat{u}_t + \psi_{wa}\hat{a}_t + \psi_{ws}\hat{s}_t.$$

Using the linear-invariance result to substitute for $\hat{\theta}_{t+\tau}$ and $\hat{w}_{t+\tau}$ in (G.12) and (G.14), I get

$$\begin{aligned}\hat{\theta}_t &= \frac{1-s}{\alpha J}\left((1-b-\psi_{wa})\hat{a}_t - \psi_{wu}\hat{u}_t - \psi_{ws}\hat{s}_t\right)\\ &+ \frac{1-s}{\alpha J}E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau\left((1-b-\psi_{wa})\hat{a}_{t+\tau} - \psi_{wu}\hat{u}_{t+\tau} - (\zeta+\psi_{ws})\hat{s}_{t+\tau}\right)\right],\end{aligned} \tag{G.16}$$

and

$$\begin{aligned}w\hat{w}_t &= \delta(1-b)\hat{a}_t + E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s)^\tau(\delta(1-b-\psi_{wa})\hat{a}_{t+\tau} - \delta\psi_{wu}\hat{u}_{t+\tau} - \delta(\zeta+\psi_{ws})\hat{s}_{t+\tau})\right]\\ &+ E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau(p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu})\hat{u}_{t+\tau}\right]\\ &+ E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau(p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa})\hat{a}_{t+\tau}\right]\\ &+ E_t\left[\sum_{\tau=1}^{\infty}\beta^\tau(1-s-p)^\tau(p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)\hat{s}_{t+\tau}\right].\end{aligned} \tag{G.17}$$

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 (G.16) and (G.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}:

$$\psi_{\theta u} = \frac{a\beta(1-s)^2 p_u}{1-a\beta(1-s)}\left(\frac{1-b-\psi_{wa}}{\alpha J}q_a - \frac{\psi_{wu}}{\alpha J}q_u - \frac{\zeta+\psi_{ws}}{\alpha J}q_s\right) - \frac{1-s}{\alpha J}\psi_{wu}, \tag{G.18}$$

$$\psi_{\theta a} = \frac{a\beta(1-s)^2 p_a}{1-a\beta(1-s)}\left(\frac{1-b-\psi_{wa}}{\alpha J}q_a - \frac{\psi_{wu}}{\alpha J}q_u - \frac{\zeta+\psi_{ws}}{\alpha J}q_s\right) + \frac{1-s}{\alpha J}(1-b-\psi_{wa}), \tag{G.19}$$

$$\psi_{\theta s} = \frac{a\beta(1-s)^2 p_s}{1-a\beta(1-s)}\left(\frac{1-b-\psi_{wa}}{\alpha J}q_a - \frac{\psi_{wu}}{\alpha J}q_u - \frac{\zeta+\psi_{ws}}{\alpha J}q_s\right) - \frac{1-s}{\alpha J}\psi_{ws}, \tag{G.20}$$

$$\begin{aligned}\psi_{wu} &= \frac{a\beta\delta(1-s)p_u}{1-a\beta(1-s)}\left[(1-b-\psi_{wa})q_a - \psi_{wu}q_u - (\zeta+\psi_{ws})q_s\right]\\ &+ \frac{a\beta(1-s-p)p_u}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta a} - (1-\delta)\psi_{wa})q_a + (p\chi(1-\alpha)\psi_{\theta u} - (1-\delta)\psi_{wu})q_u\right]\\ &+ \frac{a\beta(1-s-p)p_u}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta s} - (1-\delta)\psi_{ws} + s\chi)q_s\right],\end{aligned} \tag{G.21}$$

$$\psi_{wa} = \delta(1-b) + \frac{a\beta\delta(1-s)p_a}{1-a\beta(1-s)}\left[(1-b-\psi_{wa})q_a - \psi_{wu}q_u - (\zeta+\psi_{ws})q_s\right]$$

$$+\frac{a\beta(1-s-p)p_a}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta a}-(1-\delta)\psi_{wa})q_a+(p\chi(1-\alpha)\psi_{\theta u}-(1-\delta)\psi_{wu})q_u\right]$$

$$+\frac{a\beta(1-s-p)p_a}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta s}-(1-\delta)\psi_{ws}+s\chi)q_s\right], \tag{G.22}$$

$$\psi_{ws}=\frac{a\beta\delta(1-s)p_s}{1-a\beta(1-s)}\left[(1-b-\psi_{wa})q_a-\psi_{wu}q_u-(\zeta+\psi_{ws})q_s\right]$$

$$+\frac{a\beta(1-s-p)p_s}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta a}-(1-\delta)\psi_{wa})q_a+(p\chi(1-\alpha)\psi_{\theta u}-(1-\delta)\psi_{wu})q_u\right]$$

$$+\frac{a\beta(1-s-p)p_s}{1-a\beta(1-s-p)}\left[(p\chi(1-\alpha)\psi_{\theta s}-(1-\delta)\psi_{ws}+s\chi)q_s\right]. \tag{G.23}$$

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 (17) and (G.15) can be written in vector form as

$$f_t=\mathbb{F}(\psi_{\theta u},\psi_{\theta a},\psi_{\theta s})f_{t-1}+\epsilon_t. \tag{G.24}$$

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 (G.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 (G.18)–(G.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 (G.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 (G.18)–(G.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.