Review of Economic Studies (2023) **00**, 1–46 doi:10.1093/restud/rdad073

Advance access publication 10 July 2023

Model Complexity, Expectations, and Asset Prices

POOYA MOLAVI
Northwestern University, USA
ALIREZA TAHBAZ-SALEHI
Northwestern University, USA and CEPR, UK
and
ANDREA VEDOLIN
Boston University, NBER, USA, and CEPR, UK

First version received January 2021*; Editorial decision March* 2023*; Accepted July* 2023 *(Eds.)*

This paper analyses how limits to the complexity of statistical models used by market participants can shape asset prices. We consider an economy in which the stochastic process that governs the evolution of economic variables may not have a simple representation, and yet, agents are only capable of entertaining statistical models with a certain level of complexity. As a result, they may end up with a lower-dimensional approximation that does not fully capture the intertemporal complexity of the true data-generating process. We first characterize the implications of the resulting departure from rational expectations and relate the extent of return and forecast-error predictability at various horizons to the complexity of agents' models and the statistical properties of the underlying process. We then apply our framework to study violations of uncovered interest rate parity in foreign exchange markets. We find that constraints on the complexity of agents' models can generate return predictability patterns that are simultaneously consistent with the well-known forward discount and predictability reversal puzzles.

Key words: Subjective expectations, Model complexity, Uncovered interest rate parity

JEL codes: D84, F31, G12, G41

1. INTRODUCTION

The rational expectations framework maintains that agents have a complete understanding of their economic environment: they know the structural equations that govern the relationship between endogenous and exogenous variables, have full knowledge of the stochastic processes that determine the evolution of shocks, and are capable of forming and updating beliefs about as many variables as necessary. These assumptions are imposed irrespective of how complex the actual environment is. However, in reality, limits to cognitive and computational abilities mean that market participants are bound to rely on simplified models that may not fully account for the complexity of their environment. Thus, to the extent that agents employ such simplified models, their decisions—and any outcome that depends on those decisions—would depart from what is predicted by rational expectations.

The editor in charge of this paper was Nicola Gennaioli.

In this paper, we study how limits to the complexity of statistical models used by market participants shape asset prices. We consider a framework in which the stochastic process that governs the evolution of economic variables may not have a simple representation, and yet, agents are only capable of entertaining statistical models with a certain level of complexity. As a result, they may end up with a lower-dimensional approximation to the underlying data-generating process that does not capture its full intertemporal complexity. We show that this form of model misspecification generates systematic deviations from rational expectations with sharp predictions for the extent and nature of return and forecast-error predictability at different horizons.[1]

We present our results in the context of a simple asset pricing environment, in which a sequence of exogenous fundamentals (say, an asset's dividends) are generated by a stochastic process that can be represented as an n-factor model. While agents can observe the sequence of realized fundamentals, they neither observe nor know the underlying factors that drive them. As a result, they rely on their past observations to estimate a hidden-factor model that would allow them to make predictions about the future. As our main assumption, we assume that agents can only hold and update beliefs about models with at most k factors, where k may be distinct from the true number of factors, n. As in Molavi (2022), this assumption captures the idea that there is a limit to the complexity of statistical models that agents are able to consider, with a larger k corresponding to a more sophisticated agent who can entertain a richer class of models.

Formally, we assume that agents start with a prior belief with support over the set of k-factor models and update their beliefs over time in a Bayesian fashion. While simple, this formulation has three important features. First, the restriction on the support of agents' priors reflects our behavioural assumption that agents are incapable of holding and updating beliefs about models that are more complex than what they can entertain. Second, the assumption that agents in our framework are Bayesian ensures that any deviation from the predictions of the rational expectations benchmark is entirely due to the complexity constraint on the agents' models (as captured by the wedge between k and n). Third, this formulation implies that all parameters of the models agents use for forecasting are endogenous outcomes of learning: once we specify the number of factors k, there are no more degrees of freedom on how agents form their subjective expectations. As a result, agents' forecasts (and the resulting price, return, and forecast-error dynamics) are fully pinned down by (i) the constraint on the complexity of the agents' models and (ii) the statistical properties of the true data-generating process.

With our behavioural framework in hand, we first characterize the agents' subjective expectations in terms of the primitives of the economy. We establish that, as agents accumulate more observations, their posterior beliefs concentrate on the subset of k-factor models with minimum Kullback–Leibler (KL) divergence from the true data-generating process. This result implies that when agents can contemplate models that are as complex as the true model (that is, when $k \geq n$), they can forecast the future realizations of the fundamental as if they knew the true data-generating process. In other words, when $k \geq n$, our framework reduces to the rational expectations benchmark. Furthermore, the characterization of agents' subjective expectations in terms of the KL divergence from the true process enables us to determine price and return dynamics in the more interesting case when $k < n$.

1. While we focus on limits to the complexity of agents' statistical models, one can also consider other dimensions along which decision makers are constrained, such as limited memory (da Silveira *et al.*, 2020; Nagel and Xu, 2022), limited capacity for processing information (as in models of rational inattention), an incomplete understanding of general equilibrium effects (*e.g.* due to level-k reasoning as in Farhi and Werning, 2019 or limited "depth of knowledge" as in Angeletos and Sastry, 2021), and restrictions on the number of variables to pay attention to (Gabaix, 2014).

Having characterized agents' subjective expectations, we then turn to the main focus of our analysis: how the constraint on the complexity of statistical models used by agents shapes asset prices. More specifically, we study the extent to which the disparity between the number of factors in the agents' models (k) and that of the true data-generating process (n) can result in return and forecast-error predictability. We measure the extent of return and forecast-error predictability by relying on two families of regressions. These regressions measure the extent to which current fundamentals predict future returns and forecast errors at different horizons. Our main theoretical results provide a characterization of the slope coefficients of the aforementioned regressions in terms of the complexity of agents' models and the auto-correlation function (ACF) of the process that drives the fundamentals. Since all parameters in the agents' models, other than the number of factors k, are endogenously determined, our results provide sharp predictions for the extent of return and forecast-error predictability at different horizons.

We then use our characterization theorem to obtain a series of results that relate the term structure of the slope coefficients of the predictability regressions to the primitives of the environment. First, we show that returns and forecast errors are predictable if (and only if) the true process is more complex than the set of models considered by the agents (that is, $k < n$). Next, we show that if agents can entertain more complex models with richer time-series dependencies, their forecast errors (averaged over different horizons) become less predictable. We also obtain a characterization of the sign pattern of the slope coefficients of the return predictability regression in terms of the spectrum of the matrix that governs the evolution of the economy's fundamental. Specifically, we show that the horizons at which the slope coefficients of the return predictability regression change signs are tightly linked to this spectrum. We conclude our theoretical analysis by showing how one can back out the number of factors in the agents' model from their forecasts and the realizations of the fundamental.

We then apply our framework to study violations of uncovered interest rate parity (UIP) in foreign exchange markets. We show that the behavioural constraint on the complexity of agents' model can generate return predictability patterns that are simultaneously consistent with two well-known, but seemingly contradictory violations of UIP, namely, the forward discount and the predictability reversal puzzles. The forward discount puzzle, which dates back to Fama (1984), is the robust empirical finding that, in short time horizons ranging from a week to a quarter, high interest rate currencies tend to have positive excess returns. The predictability reversal puzzle, more recently documented by Bacchetta and van Wincoop (2010) and Engel (2016), refers to the fact that high interest rate currencies tend to have negative excess returns over longer horizons, that is, the violation of UIP reverses sign after some point. The seemingly contradictory implications of these puzzles for the relationship between currency excess returns and interest rate differentials has led some to argue for the inadequacy of existing models for explaining UIP violations (Engel, 2016).

To test our model's implications for UIP violations at different horizons, we follow Engel (2016) and run predictive regressions from a trade-weighted average currency return on the corresponding interest rate differential. We then compare the resulting slope coefficients to those implied by our theoretical framework when agents are constrained in the number of factors in their models. We find that when investors are constrained to using single-factor models, the model-implied slope coefficients line up with the ones from the data: at short horizons, deviations from rational expectations generate UIP violations that imply positive excess returns for high interest rate currencies, whereas at longer horizons the pattern reverses, with high interest rate currencies earning a negative excess return. In other words, our framework generates return predictability patterns that are simultaneously consistent with the forward discount and predictability reversal puzzles. Crucially, the model-implied return predictability pattern matches its empirical counterpart without using data on exchange rates or excess returns: the model-implied slope coefficients are constructed solely from the autocorrelation of interest rate differentials. We also find that as we increase the number of factors in agents' model, UIP violations mostly disappear at all horizons.

In the second half of our empirical analysis, we investigate whether our behavioural model's predictions are also consistent with evidence from survey data. Using surveys of professional forecasters, we estimate the extent to which past interest rate differentials predict realized forecast errors and compare the results to the extent of forecast-error predictability implied by our model. Once again, we find that, under the assumption that agents rely on single-factor models, the model-implied slope coefficients of forecast-error predictability regressions track their empirical counterparts fairly closely. In other words, the assumption that market participants may rely on simpler representations of the interest rate differential process not only leads to predictable forecast errors but also generates forecast errors that are in line with survey data. This is despite the fact that the model-implied regression coefficients are constructed solely from the autocorrelation of interest rate differentials, without relying on survey data.

1.1. *Related literature*

Our paper contributes to the literature that studies the asset pricing implications of deviations from rational expectations. The most related branch of this literature considers agents who have a misspecified view of the true data-generating process as a result of behavioural biases. For instance, Barberis *et al.* (1998) assume agents mistakenly believe the innovations in earnings are drawn from a regime with excess reversals or excess streaks, Gourinchas and Tornell (2004) consider a model in which agents misperceive the relative importance of transitory and persistent shocks, and Rabin and Vayanos (2010) show that gambler's fallacy—the belief that random sequences should exhibit systematic reversals, even in small samples—can generate momentum and reversal in returns.[2] As in these papers, deviations from rational expectations in our framework are due to misspecification in agents' models. However, we depart from the prior literature along two dimensions. First, we assume that—subject to the constraint on the complexity of their model—agents can estimate a fully flexible linear factor model with no a priori restrictions on its parameters. As a result, any distortion in agents' subjective expectations is purely due to the mismatch between the dimensions of their model and that of the true process. Second, we show that such a mismatch not only can result in predictable forecast errors that are in line with survey data but can also generate UIP violations that are simultaneously consistent with the forward discount and predictability reversal puzzles.

Another strand of the literature studies how deviations from Bayesian updating shape asset prices. The extrapolative expectations models of Hirshleifer *et al.* (2015) and Barberis *et al.* (2015) and diagnostic expectations models of Bordalo *et al.* (2018) are examples of such non-Bayesian models of updating. In contrast to this literature, we maintain the assumption of Bayesian updating and instead formulate our departure from rational expectations by assuming that agents assign zero prior beliefs to complex models with a large number of factors. This formulation allows us to impose constraints on the complexity of agents' models, while preserving other features of the rational expectations framework (such as the internal consistency of agents' subjective expectations).

2. For theoretical and empirical studies of departures from the full information rational expectations benchmark in the macroeconomics literature, see Coibion and Gorodnichenko (2015), Bordalo *et al.* (2020), Angeletos *et al.* (2020), Kohlhas and Walther (2021), and Farmer *et al.* (2023).

Our paper is also related to the broader literature that studies the implications of various kinds of constraints on agents' cognitive and computational abilities, such as imperfect or selective memory (Bordalo *et al.*, 2020; Nagel and Xu, 2022; Wachter and Kahana, 2023), limited capacity for processing information (Sims, 2003), incomplete understanding of general equilibrium effects (*e.g.* due to level-k reasoning), and restrictions on the number of variables to pay attention to (Gabaix, 2014). In contrast to these papers, agents in our framework observe the exact realization of all economic variables and have a complete understanding of the relationship between endogenous and exogenous variables. Instead, they are constrained in the complexity of the time-series models they fit to the dynamic evolution of the fundamental.

Our behavioural framework is closely related to Fuster *et al.* (2010, 2012), who study an economy in which decision makers forecast a time series using only its last k realizations and find that when the true process has hump-shaped dynamics, the misspecification in the order of the autoregressive process generates excess returns that are negatively predicted by lagged excess return. We extend this framework along two dimensions. First, rather than restricting agents to forecast a time series using its last k realizations, we allow them to estimate a fully flexible k-factor model. This is equivalent to using any arbitrary k statics of the past realizations to forecast the future.[3] Second, we use our framework to characterize the extent of return and forecast-error predictability at various horizons as a function of the number of factors k in agents' model and the statistical properties of the underlying data-generating process. Also closely related is Molavi (2022), who shows that restrictions on the dimensionality of agents' models can generate persistence bias, whereby agents attend to the most persistent observables at the expense of less persistent ones, thus inducing comovement between their various actions. We instead focus on how constraints on agents' model complexity shape the extent and nature of return and forecast-error predictability in asset pricing applications.[4]

Finally, we also contribute to the literature that studies UIP violations and in particular, the reversal in currency return predictability (Engel, 2016; Valchev, 2020; Bacchetta and van Wincoop, 2021). Within this literature, our paper is most closely related to Gourinchas and Tornell (2004), Candian and Leo (2023), and Granziera and Sihvonen (2022), who study how distortions in agents' subjective expectations can explain various exchange rate puzzles.[5] Different from these papers, which specify exogenous processes for agents' subjective expectations, the wedge between objective and subjective expectations in our framework is an endogenous outcome of learning. A consequence of this endogeneity is that, in contrast to the literature that hardwires over- and/or under-reaction of expectations into agents' models, the forecasts of agents in our framework may exhibit over- or under-reaction to news depending on the statistical properties of the underlying data-generating process.

3. We formalize this equivalence in Proposition A.1. The family of k-factor models includes all ARMA(p, q) processes such that $\max\{p, q+1\} \leq k$ and thus nests the family of AR(k) processes as a special case.

4. Methodologically, our paper is also related to the literature on model order reduction in control theory, which is concerned with characterizing lower-order approximations to large-scale dynamical systems. See Antoulas (2005) and Sandberg (2019) for textbook treatments of the subject. In contrast to this literature, which mostly focuses on optimal Hankel-norm approximations, Bayesian learning in our framework implies that agents end up with a model that minimizes the KL divergence from the true data-generating process.

5. The expectations-based violations of UIP in our framework is also related to Froot and Frankel (1989) and Chinn and Frankel (2020), who argue that most of the forward premium bias can be attributed to errors in agents' subjective expectations, as opposed to risk premia.

1.2. *Outline*

The rest of the paper is organized as follows. Section 2 presents the environment and specifies our behavioural assumption. In Section 3, we introduce the predictability regressions that we use to quantify the extent of return and forecast-error predictability. We present our main results in Section 4, where we illustrate how the behavioural constraint on the complexity of agents' models shapes return and forecast-error predictability at different horizons. In Section 5, we apply our framework to study violations of UIP in foreign exchange markets. All proofs and some additional results are provided in the Appendix. The Supplementary Appendix contains additional empirical results, as well as the proofs not presented in the main paper.

2. MODEL

We start by presenting a reduced-form asset pricing framework that serves as the basis of our analysis. This framework allows us to focus on the main ingredients of our model, while abstracting from details that are not central to the analysis. In Section 5, we provide a concrete application that can be cast into this reduced-form framework in a straightforward manner.

2.1. *Reduced-form framework*

Consider a discrete-time economy consisting of a unit mass of identical agents. Agents form subjective expectations about an exogenous sequence of variables, $\{x_t\}_{t=-\infty}^{\infty}$, which we refer to as the economy's *fundamentals*. Depending on the context, the fundamental may correspond to an asset's dividend stream over time, interest rate differential between two countries, or any sequence of variables that can be treated as exogenous in that specific application.

The sequence of fundamentals $\{x_t\}_{t=-\infty}^{\infty}$ is generated by a stationary n-factor model given by

$$\begin{aligned} z_t &= \mathbf{A}^* z_{t-1} + \mathbf{B}^* \varepsilon_t \\ x_t &= c^{*\prime} z_t, \end{aligned} \tag{1}$$

where $z_t \in \mathbb{R}^n$ denotes the vector of factors that drive the dynamic of fundamentals, $c^* \in \mathbb{R}^n$ is a vector of constants that captures the fundamental's loading on each of the n factors, and $\mathbf{A}^*$ and $\mathbf{B}^*$ are square matrices that govern the evolution of factors over time. To ensure stationarity, we assume that all eigenvalues of $\mathbf{A}^*$ are inside the unit circle: $\rho(\mathbf{A}^*) < 1$, where $\rho(\cdot)$ denotes the spectral radius. The noise terms $\varepsilon_t \in \mathbb{R}^n$ are independent over time and across factors and are normally distributed with mean zero and unit variance, that is, $\varepsilon_{it} \sim \mathcal{N}(0, 1)$. Without much loss of generality, we assume that matrix $\mathbf{B}^*$ that determines the covariance matrix of factor innovations has full rank. The process that generates the fundamentals can thus be summarized by the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$. Note that the formulation in equation (1) is flexible enough to nest the family of stationary and invertible ARMA processes (Aoki, 1983).

Throughout, we assume that there are no redundant factors in equation (1), in the sense that the dynamics of the fundamental cannot be represented by a factor model with fewer than n factors. The number of factors n thus measures the complexity of the data-generating process (1), as a larger n means that describing the dynamics of the fundamental requires a system of equations of a higher dimension.

Whereas the fundamental sequence $\{x_t\}_{t=-\infty}^{\infty}$ is assumed to be exogenous, we are interested in how the fundamentals and the agents' subjective expectations about them jointly determine an endogenous sequence of variables $\{y_t\}_{t=-\infty}^{\infty}$, which we refer to as *prices*. More specifically, we assume that the price at time t satisfies the recursive equation

$$y_t = x_t + \delta \mathbb{E}_t[y_{t+1}], \tag{2}$$

where $\mathbb{E}_t[\cdot]$ denotes the agents' time t subjective expectation and $\delta \in (0, 1]$ is a constant. Equation (2) is akin to a standard no-arbitrage condition with a natural interpretation. For example, if the sequence $\{x_t\}_{t=-\infty}^{\infty}$ represents an asset's dividend stream, then y_t corresponds to the price of the asset at time t, with equation (2) simply capturing the fact that the asset's price is the sum of its dividends at that time and its expected future price, discounted at some rate δ. We also note that agents' subjective expectations, $\mathbb{E}_t[\cdot]$, may be distinct from the expectations arising from the true data-generating process (1), which we denote by $\mathbb{E}_t^*[\cdot]$.

Given the sequence of fundamentals and prices $\{(x_t, y_t)\}_{t=-\infty}^{\infty}$, we define *excess returns* at time $t+1$ as the sum of the fundamental and the change in price between t and $t+1$, properly discounted:

$$\mathrm{rx}_{t+1} = \delta y_{t+1} - y_t + x_t. \tag{3}$$

Using equation (2), we can also express excess returns as the difference between the realized and the expected price: $\mathrm{rx}_{t+1} = \delta(y_{t+1} - \mathbb{E}_t[y_{t+1}])$.

Together with the specification of how agents' subjective expectations are formed, equations (1)–(3) fully describe our environment. Note that while we have abstracted from certain details of the economy (such as preferences, endowments, and the market structure), our subsequent analysis applies to any model with the same reduced-form representation as the above framework. As already mentioned, in Section 5, we provide a concrete application that can be cast into this reduced-form framework in a straightforward manner.

2.2. *Constraints on model complexity*

While at any given time t, agents observe the sequence of realized fundamentals up to that time, we assume that they observe neither the underlying factors $(z_1, \ldots, z_n)$ nor the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$ that governs the data-generating process. As a result, they use the past realizations of the fundamental to learn about the underlying process that generates $\{x_t\}_{t=-\infty}^{\infty}$, which they then use to make forecasts about the future.

As our main behavioural assumption, we assume that agents face a constraint on the complexity of statistical models they can consider. Specifically, as in Molavi (2022), we assume that they can only entertain models with at most k factors, where k may be distinct from the true number of factors, n, in equation (1). The number of factors, k, thus indexes the agents' degree of sophistication, with a larger k corresponding to agents who can entertain a richer class of models.

Formally, we assume that agents can only hold and update beliefs over the set of k-factor models $\Theta_k = \{(\mathbf{A}, \mathbf{B}, c) : \mathbf{A}, \mathbf{B} \in \mathbb{R}^{k \times k}, \rho(\mathbf{A}) < 1, \text{ and } c \in \mathbb{R}^k\}$ of the following form:

$$\begin{aligned} \omega_t &= \mathbf{A}\omega_{t-1} + \mathbf{B}\varepsilon_t \\ x_t &= c'\omega_t, \end{aligned} \tag{4}$$

where $\omega_t \in \mathbb{R}^k$ denotes the vector of k underlying factors, $(\mathbf{A}, \mathbf{B}, c) \in \Theta_k$ parameterizes the process that govern the factors' evolution and the fundamental's loading on each of the factors, and the noise terms $\varepsilon_{it} \sim \mathcal{N}(0, 1)$ are independent over time and across factors. Unless otherwise stated—and to simplify the analysis—we assume that $\mathbf{B}$ has full rank. Note that, with some abuse of notation, we can write $\Theta_k \subseteq \Theta_{k+1}$, thus capturing the fact that agents with a higher k can contemplate a larger class of models.

A few remarks are in order. First, note that we impose no restrictions, other than the number of factors, on the agents' models: the k factors $(\omega_1, \ldots, \omega_k)$ in the agents' model may overlap with a subset of the n factors $(z_1, \ldots, z_n)$ that drive the fundamental, may be linear combinations of the underlying n factors, or can be constructed in an entirely different way altogether. Second, as we discuss in further detail below, the k-factor model agents use for forecasting is an endogenous outcome of learning: agents will rule out models that are less consistent with their past observations in favour of those that are more so (in a sense that we will formalise). Third, when $k < n$, the set of models entertained by agents does not contain the true n-factor data-generating process in equation (1), that is, $\theta^* \notin \Theta_k$. In such a case, our behavioural assumption implies that irrespective of which k-factor model they use, agents will end up with a misspecified model of the world. This observation also clarifies the bite of our behavioural assumption: whereas more sophisticated agents with $k \geq n$ can recover the model that generates the fundamentals (at least in principle), those with $k < n$ can at best construct lower-dimensional approximations to the true data-generating process.

2.3. *Subjective expectations and learning*

Agents form their subjective expectations by learning from the past realizations of the fundamental and updating their beliefs in a Bayesian fashion. Specifically, we assume that they start with a common prior belief, $\mu_{t_0} = \bar{\mu} \in \Delta\Theta_k$, with support over an arbitrary compact subset of the set of k-factor models at some initial period t_0 and form Bayesian posterior beliefs $\mu_t \in \Delta\Theta_k$ after observing the sequence of fundamentals $\{x_{t_0}, \ldots, x_{t-1}, x_t\}$.

This formulation has a few important implications. First, it implies that agents assign a zero prior belief to all models with more than k factors, thus reflecting our assumption that they are incapable of holding and updating beliefs about models that are more complex than what they can entertain. Second, aside from the constraint on the number of factors, we allow agents to estimate a fully flexible model with arbitrary factor dynamics ($\mathbf{A}$), volatility of and covariance between innovations ($\mathbf{B}$), and loading of the fundamental on each of the factors (c). Third, and most importantly, agents in our framework are fully Bayesian with an internally consistent system of beliefs, in the sense that their subjective expectations satisfy the law of iterated expectations: $\mathbb{E}_t\mathbb{E}_{t+h}[\cdot] = \mathbb{E}_t[\cdot]$ for all $h \geq 0$ and all t. This means that, under agents' subjective expectations, all forecast errors at time $t+h$ are unpredictable given their information set at time t. When coupled with equations (2) and (3), this observation also implies that excess returns in our framework are equal to the discounted sum of agents' forecast revisions:[6]

$$\mathrm{rx}_{t+1} = \sum_{\tau=1}^{\infty} \delta^{\tau} \left(\mathbb{E}_{t+1}[x_{t+\tau}] - \mathbb{E}_t[x_{t+\tau}] \right). \tag{5}$$

Throughout the rest of the paper, we restrict our attention to agents' long-run beliefs by assuming that they have already observed a long sequence of past realizations of the fundamental. Formally, we assume that at any given period t, agents observe the sequence $\{x_{t_0}, \ldots, x_{t-1}, x_t\}$, and consider the limit as $t_0 \to -\infty$. This assumption allows us to study the systematic implications of the complexity constraint on agents' models (*i.e.* the disparity between k and n), while abstracting from fluctuations in beliefs that arise due to observing a finite sample.

6. Throughout, we restrict our attention to the unique "fundamental solution" of recursive equation (2), which is the solution that does not depend on any sunspot shocks.

2.4. *Discussion*

We conclude this section with a brief discussion of our framework and its key ingredients.

As already mentioned, the behavioural restriction on the maximum dimension of agents' model captures the idea that agents may not be able to fully comprehend the time-series relationships in the data-generating process. Such a lack of comprehension can be due to constraints on cognition or memory, which may make it difficult for agents to use higher-dimensional models that are not easy to understand, use, or revise. We provide one example of such a restriction in Appendix A.1, where we show that an agent who is restricted to the class of k-factor models is isomorphic to an agent with limited recall, who makes predictions about the future by relying on at most k summary statistics of the entire history of her past observations. Under this interpretation, an agent with a larger k can construct a richer set of statistics, maintain a more complete record of the past, and hence better capture the intertemporal dependencies in the underlying data-generating process.

Despite their limited ability to comprehend the complex time-series properties of the data-generating process, agents are assumed to fully understand (and have no uncertainty about) the relationship between endogenous and exogenous variables in equation (2). This is akin to assuming that agents have a correctly specified model of the economy's structure, such as preferences, market structure, market clearing, and all partial and general equilibrium effects. The only potential source of misspecification in agents' model is in the stochastic process that drives the fundamental.[7]

The assumption that agents in our framework are Bayesian with access to a long history of observations is meant to ensure that any deviation from the predictions of textbook rational expectations is entirely due to the complexity constraint on their models. Indeed, as we show in subsequent sections, when the agents' set of models is flexible enough to contain the true data-generating process (*i.e.* when $k \geq n$), Bayesian learning guarantees that agents can forecast future realizations of the fundamental as if they knew the true process. In such a case, our framework coincides with the benchmark rational expectations framework.

Bayesian updating also implies that agents' posterior beliefs at any given time only depend on the sequence of realized fundamentals up to that time and is independent of other characteristics of the economy. Thus, holding their information sets constant, agents in our framework end up with the same posterior beliefs irrespective of how preferences, endowments, or other features of the economy are specified.[8]

Another important feature of our framework is that all parameters of the model agents use for forecasting are endogenous outcomes of learning: once we specify the number of factors k, there are no more degrees of freedom on how agents form their expectations. As a result, agents' forecasts (and the resulting price and return dynamics) are fully pinned down by (i) the limit on the complexity of their model, as captured by k, and (ii) the true data-generating process (1) as summarized by $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$. The endogeneity of agents' model also implies that the way their forecasts respond to new information cannot be decoupled from the environment they live in. In particular, as we show in the Supplementary Appendix D, agents' forecasts may exhibit systematic over- or under-reaction to news depending on the statistical properties of the

7. In view of equations (2) and (3), this assumption also implies that all realizations of prices and returns up to time t are measurable with respect to the information set generated by $\{x_\tau\}_{\tau \leq t}$. Therefore, irrespective of whether past prices and returns are observable or not, it is sufficient to consider expectations conditioned on the past realizations of the fundamental.

8. Also, note that the assumptions that agents (i) are Bayesians and (ii) assign a zero prior belief to models with more than k factors imply that no data or statistical analysis would make them revise their beliefs in such a way that they end up assigning a non-zero posterior belief to models with more than k factors.

underlying data-generating process. This is in contrast to the literature that hardwires over- and under-reaction of expectations into agents' models.

As a final remark, we note that our focus on the simple no-arbitrage condition in equation (2) abstracts from potentially important dimensions—such as time-varying discount rates—that matter for price and return dynamics. Nonetheless, this choice allows us to isolate, in the most transparent manner, how the agents' inability to entertain complex models distorts their subjective expectations and shapes return dynamics.

3. RETURN AND FORECAST-ERROR PREDICTABILITY

Under rational expectations, all future excess returns and forecast errors are unpredictable given past information. As a result, the extent to which returns and forecast errors are predictable can serve as a natural measure for how the departure from rational expectations caused by our behavioural assumption shapes asset price dynamics.

In this section, we first introduce the regressions that we use to assess the implications of the wedge between the complexity of agents' model and the true data-generating process for return and forecast-error predictability. We then characterize the slope coefficients of these predictability regressions in terms of the statistical properties of the agents' model and the true data-generating process.

3.1. *Return predictability*

We measure the extent of return predictability by relying on a family of linear regressions that quantify the extent to which the current realization of the fundamental predicts future excess returns at different horizons:

$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}} x_t + \varepsilon_{t,h}, \tag{6}$$

where $h \geq 1$ denotes the horizon and rx is the excess return, as defined in equation (3). If returns are unpredictable—as would be the case under rational expectations—then $\beta_h^{\mathrm{rx}} = 0$ for all horizons h. If, on the other hand, an increase in the fundamental today is associated with higher excess returns at horizon h, then $\beta_h^{\mathrm{rx}} > 0$. Thus, the family of slope coefficients $(\beta_1^{\mathrm{rx}}, \beta_2^{\mathrm{rx}}, \ldots)$ not only provides a measure for departures from the rational expectations benchmark but also determines the extent to which such departures vary with horizon h.[9]

3.2. *Forecast-error predictability*

We test for the predictability of forecast errors by relying on a similar family of linear regressions. These regressions, which resemble the specification in Kohlhas and Walther (2021), measure the extent to which the current realization of the fundamental predicts m-step-ahead forecast errors at a given horizon $h \geq 1$:

$$x_{t+h+m-1} - \mathbb{E}_{t+h-1}[x_{t+h+m-1}] = \alpha_{h,m}^{\mathrm{fe}} + \beta_{h,m}^{\mathrm{fe}} x_t + \varepsilon_{t,h,m}, \tag{7}$$

9. The family of regressions in equation (6) resembles Fama's (1984) regression specification for testing deviations from the UIP condition. In that context, the fundamental, x_t, corresponds to the log interest rate differential between two countries, y_t is the log of the exchange rate, and rx_t is the currency excess return at time t. See Section 5 for a more detailed discussion of the application of our framework to foreign exchange markets.

where $\mathbb{E}[\cdot]$ denotes the agents' subjective expectations. The unpredictability of forecast errors under rational expectations requires that $\beta^{\text{fe}}_{h,m} = 0$ for all $h, m \geq 1$. Therefore, once again, the term structure of coefficients $\beta^{\text{fe}}_{h,m}$ provides us with a natural measure for assessing the implications of our behavioural framework at different time horizons relative to the benchmark of rational expectations.

3.3. *Characterization*

In the remainder of this section, we characterize the slope coefficients of regressions (6) and (7) in terms of the statistical properties of the agents' model and the true data-generating process.

Recall that the key friction in our framework is that agents rely on potentially simplified models that do not capture the full intertemporal complexity of the true data-generating process. This means that they may misperceive the autocorrelation of the process that generates the fundamental. To formalize this misperception, let $\xi^*_\tau = \mathbb{E}^*[x_t x_{t+\tau}]/\mathbb{E}^*[x_t^2]$ denote the autocorrelation of the fundamental at lag τ under the true data-generating process. Similarly, let $\xi_\tau = \mathbb{E}[x_t x_{t+\tau}]/\mathbb{E}[x_t^2]$ denote the same autocorrelation under agents' subjective expectations, which may be distinct from ξ^*_τ. The following result expresses the large-sample limits of the slope coefficients of regressions (6) and (7) in terms of these autocorrelations.

Proposition 1. *The slope coefficients of the return and forecast-error predictability regressions* (6) *and* (7) *are given by*

$$\beta^{\text{rx}}_h = \delta \frac{(\xi^*_h - \xi_h) - \sum_{\tau,s=1}^{\infty} \Xi^{-1}_{\tau s} \xi_s (\xi^*_{h-\tau} - \xi_{h-\tau})}{1 - \sum_{\tau,s=1}^{\infty} \delta^\tau \, \Xi^{-1}_{\tau s} \xi_s} \tag{8}$$

and

$$\beta^{\text{fe}}_{h,m} = (\xi^*_{h+m-1} - \xi_{h+m-1}) - \sum_{\tau=1}^{\infty} \sum_{s=1}^{\infty} \xi_{s+m-1} \, \Xi^{-1}_{\tau s} (\xi^*_{h-\tau} - \xi_{h-\tau}), \tag{9}$$

respectively, where Ξ *is an infinite-dimensional matrix such that* $\Xi_{ij} = \xi_{i-j}$ *for all* $i, j \geq 1$.

This result has a few immediate implications. First, it is straightforward to verify that, if $\mathbb{E}[\cdot] = \mathbb{E}^*[\cdot]$, then $\xi_\tau = \xi^*_\tau$ for all τ and hence, $\beta^{\text{rx}}_h = \beta^{\text{fe}}_{h,m} = 0$ at all horizons $h \geq 1$. Thus, when agents use the correct model to make forecasts about the future realizations of the fundamental—as would be the case with rational expectations—neither excess returns nor forecast errors are predictable given past realizations of the fundamental. Second, when agents' subjective expectations do not coincide with expectations under the true data-generating process (so that $\xi_h \neq \xi^*_h$ for some h), the coefficients of the predictability regressions may in general be different from zero. This is despite the fact that agents' expectations are internally consistent and satisfy the law of iterated expectations. Finally, Proposition 1 illustrates that the wedge between the agents' subjective expectations and rational expectations may have differential impacts on the slope coefficients at different horizons h.

Taken together, equations (8) and (9) illustrate how the statistical properties of the true process and agents' subjective model—as summarized by the corresponding autocorrelations—are sufficient statistics for the extent of return and forecast-error predictability. In fact, according to these equations, the term structures of the coefficients of the predictability regressions are closely linked to whether agents over- or under-estimate the true autocorrelation of the data-generating process at different horizons (that is, whether $\xi^*_\tau < \xi_\tau$ or $\xi^*_\tau > \xi_\tau$). However, recall from Section 2 that agents' subjective expectations are themselves endogenous outcomes of learning. This means that, in our framework, the autocorrelation of the fundamental perceived by the agents, $\{\xi_\tau\}_{\tau=-\infty}^{\infty}$, is also endogenous. Therefore, to apply Proposition 1 to our framework, we first need to characterize agents' subjective expectations in terms of the primitives of the economy, namely, (i) the true data-generating process and (ii) the complexity of the models agents can entertain, as summarized by the maximum number of factors in their model, k.

4. MODEL COMPLEXITY AND ASSET PRICES

With the characterization result in Proposition 1 in hand, we now turn to the main focus of our analysis: how the constraint on the complexity of statistical models used by market participants can shape asset prices. More specifically, we study the extent to which the disparity between the number of factors in agents' models (k) and that of the true data-generating process (n) determines the statistical properties of the agents' models and hence can result in (excess) return and forecast-error predictability.

4.1. *Subjective expectations under complexity constraint*

As a first step, we characterize agents' subjective expectations in terms of model primitives.

Recall from Section 2 that the true data-generating process can be represented by the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$, all of which are of dimension n, whereas agents' models are restricted to the set of k-dimensional models, Θ_k. Given an arbitrary model $\theta \in \Theta_k$ for the agents, denote the KL divergence of model θ from the true process by

$$\mathrm{KL}(\theta^* \| \theta) = \mathbb{E}^*[-\log f^\theta(x_{t+1}|x_t, \ldots)] - \mathbb{E}^*[-\log f^*(x_{t+1}|x_t, \ldots)], \tag{10}$$

where f^* is the density of the fundamental under the true data-generating process, f^θ is the agents' subjective density under model θ, and $\mathbb{E}^*[\cdot]$ is the expectation with respect to the true process. It is well known that the KL divergence is always non-negative and obtains its minimum value of zero if and only if densities f^θ and f^* coincide almost everywhere. As such, $\mathrm{KL}(\theta^* \| \theta)$ can be interpreted as a measure for the disparity between agents' subjective expectations under model θ and the true process.

The next result characterizes agents' subjective expectations in terms of the KL divergence from the true data-generating process. As discussed in Subsection 2.3, we focus on the outcome of learning under the assumption that agents have access to a long history of the past realizations of the fundamental. This allows us to study the systematic implications of the complexity constraint for agents' subjective expectations, while abstracting from the fluctuations in beliefs that arise due to a finite sample.

Proposition 2. *Let $\widehat{\Theta}_k \subseteq \Theta_k$ be an arbitrary compact set with $\bar{\mu}(\widehat{\Theta}_k) = 1$ and let $\tilde{\Theta} \subseteq \widehat{\Theta}_k$ denote a collection of k-factor models with a positive measure under the agents' prior beliefs, i.e. $\bar{\mu}(\tilde{\Theta}) > 0$. If*

$$\operatorname*{ess\,inf}_{\theta \in \tilde{\Theta}} \mathrm{KL}(\theta^* \| \theta) > \min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta),$$

then

$$\lim_{t_0 \to -\infty} \mu_t(\tilde{\Theta}) = 0 \quad \mathbb{P}^*\textit{-almost surely}.$$

This result, which is in line with the literature on learning under model misspecification (such as Berk, 1966; Esponda and Pouzo, 2016), states that agents' posterior beliefs concentrate on the subset of k-factor models that have minimum KL divergence from the true data-generating process. When agents can contemplate models that are as complex as the true model (that is, when

$k \geq n$, and as a result, $\theta^* \in \Theta_k$), Proposition 2 implies that they assign a posterior belief of zero to all models with a non-zero KL divergence from the true process θ^*. In other words, agents can forecast the future realizations of the fundamental as if they knew the true process (*i.e.* $\mathbb{E}[\cdot] = \mathbb{E}^*[\cdot]$), in which case our framework coincides with the benchmark of rational expectations. In contrast, when the model complexity constraint binds (*i.e.* when $k < n$ so that $\theta^* \notin \Theta_k$), Proposition 2 implies that agents become convinced, mistakenly, that the fundamental is generated according to the stochastic process that is closest to the true process as measured by KL divergence. In juxtaposition with Proposition 1, this result provides us with the necessary ingredients to characterize the implications of the behavioural constraint on the complexity of agents' models for return and forecast-error predictability.

4.2. *Return and forecast-error predictability*

We now use Propositions 1 and 2 to characterize how the constraint on the complexity of agents' models shapes the extent of return and forecast-error predictability, as captured by the term structures of slope coefficients of regressions (6) and (7). As already discussed, all parameters of the model agents use for forecasting are endogenous outcomes of learning. As a result, once we specify the maximum number of factors k in the agents' model, there are no more degrees of freedom on how they form their expectations. This, in turn, implies that the autocorrelation of the fundamental as perceived by the agents and hence the coefficients of return and forecast-error predictability regressions can be expressed only in terms of the number of factors k in the agents' model and the statistical properties of the true data-generating process. We have the following result.

Theorem 1. *Suppose agents are constrained to k-factor models and let* $\{\xi^*_\tau\}_{\tau=-\infty}^{\infty}$ *denote the true autocorrelation of the fundamental.*

(1) The autocorrelation of the fundamental according to agents' subjective expectations is

$$\xi_h = \frac{\sum_{s=0}^{\infty}(u'\mathbf{M}^{h+s}u)(u'\mathbf{M}^s u)}{\sum_{s=0}^{\infty}(u'\mathbf{M}^s u)^2} \quad \text{for all } h \geq 0, \tag{11}$$

where $\mathbf{M}$ *and* u *are, respectively, a* $k \times k$ *stable matrix and a* $k \times 1$ *unit vector that minimize*

$$H(\mathbf{M}, u) = 1 - 2\sum_{s=1}^{\infty}\phi_s^{(1)}\xi_s^* + \sum_{s=1}^{\infty}\sum_{\tau=1}^{\infty}\phi_s^{(1)}\phi_\tau^{(1)}\xi_{\tau-s}^* \tag{12}$$

subject to $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}') < 1$ *and* $\phi_s^{(m)} = u'\mathbf{M}^{m-1}[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$.

(2) The coefficients of the return and forecast-error predictability regressions are given by

$$\beta_h^{\mathrm{rx}} = \delta\frac{\xi_h^* - \sum_{\tau=1}^{\infty}\phi_\tau^{(1)}\xi_{h-\tau}^*}{1-\sum_{\tau=1}^{\infty}\delta^\tau\phi_\tau^{(1)}} \tag{13}$$

$$\beta_{h,m}^{\mathrm{fe}} = \xi_{h+m-1}^* - \sum_{\tau=1}^{\infty}\phi_\tau^{(m)}\xi_{h-\tau}^*. \tag{14}$$

This theorem is the main characterization result of the paper. It relates the autocorrelation of the fundamental as perceived by agents and the resulting coefficients of predictability regressions to (i) the number of factors in agents' model and (ii) the statistical properties of the true data-generating process. The theorem follows from combining Proposition 1 with Proposition 2, according to which agents' beliefs concentrate on the subset of k-factor models with minimum

KL divergence from the true data-generating process. Even though the objective function in equation (12) is in terms of matrix $\mathbf{M}$ and vector u, these objects are closely linked to the parameters of agents' model, $\theta = (\mathbf{A}, \mathbf{B}, c)$.[10]

The characterization result in Theorem 1 serves two distinct purposes in our analysis. First, it allows us to directly apply our framework to asset pricing applications. In particular, in any context in which endogenous and exogenous variables are related to one another via equation (2), we can use the autocorrelation of the exogenous variables and the expressions in equations (13) and (14) to compute the implied coefficients of the predictability regressions. This is the approach we take in the next section. Second, Theorem 1 enables us to perform comparative static analyses with respect to the primitives of the economy and to compare the extent of return and forecast-error predictability at different horizons, as we do in the remainder of this section.

We start with a simple result that considers the case where agents' models are sufficiently rich to fully capture the statistical properties of the data-generating process.

Proposition 3. *Let k and n denote the number of factors in, respectively, agents' model and the true data-generating process.*

(1) If $k \geq n$, then $\beta_h^{\mathrm{rx}} = \beta_{h,m}^{\mathrm{fe}} = 0$ for all horizons h and all m.
(2) If $k < n$, then there exist h, $\tilde{h}$, and m such that $\beta_h^{\mathrm{rx}} \neq 0$ and $\beta_{\tilde{h},m}^{\mathrm{fe}} \neq 0$.

The first statement establishes that if agents can contemplate models that are as complex as the true model, then there is no return or forecast-error predictability at any horizon. This is a consequence of the fact that agents in our framework are Bayesian with access to a long history of observations. Bayesian updating, together with the assumption that $k \geq n$, implies that agents rule out models that are inconsistent with their past observations in favour of the true data-generating process, while the large sample size guarantees that they do not suffer from finite-sample problems such as overfitting. They recover the underlying data-generating process even if they use models that have too many parameters relative to the true model (*i.e.* when $k > n$). The second statement of Proposition 3 shows that the converse implication is also true: if there is no predictability, the set of models entertained by agents has to be rich enough to contain the true underlying model. This result follows from the characterization in Proposition 1, which implies that $\beta_h^{\mathrm{rx}} = \beta_{h,m}^{\mathrm{fe}} = 0$ at all horizons h if and only if the model-implied autocorrelation coincides with the true autocorrelation of the data-generating process at all lags ($\xi_\tau = \xi_\tau^*$ for all τ). But this can happen only if agents are capable of entertaining complex enough models to understand the time-series dependencies of the underlying process: when $k < n$, agents are bound to misperceive the persistence of the data-generating process for at least one horizon, which in turn manifests itself as predictable returns and forecast errors.

Our next result concerns the extent of return predictability at long horizons.

Proposition 4. *Excess returns and forecast errors are not predictable in the long run:*

$$\lim_{h\to\infty} \beta_h^{\mathrm{rx}} = \lim_{h\to\infty} \beta_{h,m}^{\mathrm{fe}} = 0.$$

This result is a direct consequence of the fact that both the true data-generating process (1) and agents' subjective model (4) are ergodic and stationary. As a result, the effect of time-t variables on returns and forecast errors at time $t+h$ die out eventually as h increases, irrespective of whether the complexity constraint binds or not.

10. For example, as we show in the proof of Theorem 1, the eigenvalues of matrix $\mathbf{M}$ that solves (12) coincide with the eigenvalues of matrix $\mathbf{A}$ that minimizes the KL divergence in equation (10).

Next, we study how changes in the number of factors in agents' model determine forecast-error predictability. Let $(\phi_1^*, \phi_2^*, \ldots)$ denote the coefficients of the autoregressive representation of the true data-generating process: $x_t = \sum_{\tau=1}^{\infty} \phi_\tau^* x_{t-\tau} + \varepsilon_t$ for some independent and identically distributed process ε_t.[11] Note that these coefficients only depend on the statistical properties of the true data-generating process and are independent of agents' subjective model (or its complexity). We have the following result.

Proposition 5. *Let $\bar{\beta}^{\text{fe}} = \sum_{h=1}^{\infty} \phi_h^* \beta_{h,1}^{\text{fe}}$, where $\beta_{h,m}^{\text{fe}}$ is the coefficient of the forecast-error predictability regression (7). Then,*

(1) $\bar{\beta}^{\text{fe}} > 0$ for all $k < n$;
(2) $\bar{\beta}^{\text{fe}}$ is decreasing in k;
(3) $\bar{\beta}^{\text{fe}} = 0$ for all $k \geq n$.

This comparative static result captures the intuitive idea that if agents can entertain more complex models with richer time-series dependencies, their forecast errors (averaged over different horizons) become less predictable. More specifically, Proposition 5 establishes that, holding the underlying data-generating process fixed, a particular discounted sum of the term structure of the coefficients of the predictability regression (7) is strictly positive for small values of k, decreases as the agents rely on more complex models, and becomes equal to zero when k exceeds n.

4.3. *Single-factor models*

While the optimization problem in Theorem 1—which relates return and forecast-error predictability to the primitives of the economy—does not have a closed-form characterization for a general k, it is possible to obtain such a characterization for the case in which agents are restricted to the class of single-factor models. Despite being only a special case, this closed-form characterization is a transparent and easy-to-use result that is informative about how the extent of return and forecast-error predictability varies with horizon h and the statistical properties of the true data-generating process.

Proposition 6. *If agents are constrained to single-factor models, then the coefficients of the return and forecast-error predictability regressions are given by*

$$\beta_h^{\text{rx}} = \frac{\delta}{1 - \delta\xi_1^*}(\xi_h^* - \xi_{h-1}^*\xi_1^*) \tag{15}$$

$$\beta_{h,m}^{\text{fe}} = \xi_{h+m-1}^* - \xi_1^{*m}\xi_{h-1}^*, \tag{16}$$

respectively, where ξ_h^* is the autocorrelation of the fundamental at lag h.

The characterization in Proposition 6 allows us to explore the determinants of the term structure of the predictability coefficients in further detail. We are particularly interested in the sign patterns of the return predictability coefficients, that is, whether higher realizations of the fundamental at time t predict higher or lower excess returns at time $t + h$. Our next result shows that the spectrum of matrix $\mathbf{A}^*$ in equation (1), which governs the evolution of the fundamental, plays a central role in shaping the signs of the coefficients of the return predictability regression.

11. Any process in the form of equation (1) has an invertible representation, and hence, can be expressed as an autoregressive process (with potentially infinitely many lags).

Proposition 7. *Suppose $k=1$ and $n=2$. Furthermore, suppose $\mathbf{A}^*$ has two distinct eigenvalues, λ_1 and λ_2, such that $|\lambda_1| \geq |\lambda_2|$.*

(1) If both eigenvalues are real and $\lambda_1 > 0$, then β_h^{rx} never changes sign.
(2) If both eigenvalues are real and $\lambda_1 < 0$, then β_h^{rx} changes sign at every horizon.
(3) If both eigenvalues are complex, then β_h^{rx} changes sign at horizons $h = r\pi/|\varphi|$ for $r \in \mathbb{N}$, where $\varphi = \arg(\lambda_1)$ is the argument of λ_1 in its polar representation.

That the return predictability pattern depends on the eigenvalues of $\mathbf{A}^*$ is a consequence of the fact that these eigenvalues are what determine the behaviour of the autocorrelation of the data-generating process at different horizons. For example, when $\mathbf{A}^*$ has a complex conjugate pair of eigenvalues, ξ_h^* exhibits an oscillatory behaviour. This pattern is then naturally reflected in the behaviour of β_h^{rx}, as agents who are restricted to one-factor models fit a process with an exponentially decaying autocorrelation to a time series with a more complex autocorrelation.

4.4. *Identifying the complexity of agents' model*

Our results so far were focused on characterizing the extent of return and forecast-error predictability as a function of the complexity of agents' model. We conclude this section by establishing how one can back out the number of factors in agents' model from their forecasts and the realizations of the fundamental.

Consider the following representation of agents' m-step-ahead forecasts in terms of the history of their observations:

$$\mathbb{E}_t[x_{t+m}] = \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} x_{t+1-\tau}, \tag{17}$$

where $(\phi_1^{(m)}, \phi_2^{(m)}, \ldots)$ are a collection of coefficients that are independent of the fundamental's realizations. Given these coefficients, let $\mathbf{\Phi}$ denote the (infinite-dimensional) Hankel matrix with typical element $\mathbf{\Phi}_{ij} = \phi_{i+j-1}^{(m)}$ for all $i, j \geq 1$.[12] We have the following result.

Proposition 8. *Suppose agents are restricted to the class of k factor models, where $k \leq n$. Then,*

$$\mathrm{rank}(\mathbf{\Phi}) = k \tag{18}$$

generically in $\theta \in \Theta_k$, where $\mathbf{\Phi}$ is the Hankel matrix constructed from coefficients $(\phi_1^{(m)}, \phi_2^{(m)}, \ldots)$ in equation (17).

This result establishes a key property of agents' expectations and illustrates how one can use agents' forecasts to determine the number of factors in their model. Proposition 8 follows from the particular structure of matrix $\mathbf{\Phi}$. The statement that rank($\mathbf{\Phi}$)$= k$ means that all rows of $\mathbf{\Phi}$ can be expressed as linear combinations of its first k rows, which in turn implies that coefficients $(\phi_{2k+1}^{(m)}, \phi_{2k+2}^{(m)}, \ldots)$ cannot be determined independently from $(\phi_1^{(m)}, \ldots, \phi_{2k}^{(m)})$.[13] Thus, an increase in rank($\mathbf{\Phi}$) translates into expanding the possibilities of how agents can make forecasts

12. We suppress the dependence of $\mathbf{\Phi}$ on m for notational simplicity. Matrix $\mathbf{X}$ is said to be a *Hankel matrix* if there exists a vector x such that $\mathbf{X}_{ij} = x_{i+j-1}$ for all i and j.

13. That all rows of an infinite-dimensional Hankel matrix of rank k can be written as a linear combination of its first k rows is a consequence of what is known as Kronecker's Theorem. See Theorem 2 of Al'pin (2017) for the statement.

using their past observations in the form of equation (17)—as one would anticipate to be the case when agents rely on a richer set of models with a larger number of factors, k.[14]

We conclude with two remarks about the statement of Propositions 8. First, note that one can recover the number of factors in agents' model from their forecasts only when $k \leq n$. When $k > n$, agents' forecasts coincide with forecasts under rational expectations, which means agents form the same expectations irrespective of the actual value of k. Second, note that when $k < n$, it may still be the case that the model in Θ_k that minimizes the KL divergence from the true data-generating process ends up having a lower dimension than k. However, this is a non-generic scenario, as the minimizer of equation (10) generically falls in $\Theta_k \setminus \Theta_{k-1}$.

5. APPLICATION: VIOLATIONS OF UIP

In this section, we apply our framework to study violations of UIP in foreign exchange markets and investigate the extent to which limits to the complexity of statistical models that agents use for forecasting can explain various exchange rate puzzles.

We start by providing some background and reproducing the empirical evidence on UIP violations. We then use the characterization result in Theorem 1 to test our behavioural model's predictions for the slope coefficients of return and forecast-error predictability regressions (6) and (7) in this context.

5.1. *Background*

One of the central tenets of international finance is the uncovered interest rate parity condition, which maintains that high interest rate currencies should depreciate vis-à-vis those with low interest rates. Yet—in what has become known as the "forward discount puzzle"—a vast empirical literature documents that, over short time horizons (ranging from a week to a quarter), high interest rate currencies tend to appreciate. In other words, short-term deposits of high interest rate currencies tend to earn a predictively positive excess return.

More recently, Bacchetta and van Wincoop (2010) and Engel (2016) document a distinct but related puzzle, known as the "predictability reversal puzzle." They find that UIP violations reverse sign over longer horizons, with high interest rate currencies earning negative excess returns at horizons from 4 to 7 years. The seemingly contradictory implications of the forward discount and predictability reversal puzzles for the relationship between currency excess returns and interest rate differentials has led some to argue for the inadequacy of existing models for explaining UIP violations. For example, Engel (2016) argues that risk-based explanations of the forward discount puzzle—which attribute the violations of UIP to the relative riskiness of holding short-term deposits in high interest rate countries—cannot account for the predictability reversal puzzle.

To reproduce the empirical evidence on the above-mentioned UIP violations, we first map this context to our reduced-form environment in Subsection 2.1. Let the fundamental denote the log interest rate differential between the US and a foreign country, that is, $x_t = i_t^\star - i_t$, where i_t and $i_t^\star$ are nominal interest rates on deposits held in US dollars and the foreign currency, respectively. Also, let y_t denote the log of the foreign exchange rate, expressed as the US dollar

14. The statement of Proposition 8 is closely related to the well-known result in control theory that the order of a dynamical system coincides with the rank of the system's Hankel operator. For a textbook treatment, see, for example, Antoulas (2005, Theorem 4.40). Proposition 8 generalizes and adapts this result to our behavioural framework by allowing for expectations that are formed using a misspecified model.

price of the foreign currency. With the discount rate set to $\delta = 1$, equation (2) is then nothing but the (log-linearized) interest rate parity condition:

$$y_t = (i_t^\star - i_t) + \mathbb{E}_t[y_{t+1}]. \tag{19}$$

Note that, as in Subsection 2.1, the expectation in (19) denotes investors' subjective expectations, which may differ from those arising from the true data-generating process. Finally, equation (3) is simply the definition of currency excess returns:

$$\mathrm{rx}_{t+1} = y_{t+1} - y_t + (i_t^\star - i_t).$$

We reproduce the empirical findings on UIP violations at different horizons by following Engel (2016) and building a trade-weighted average exchange rate and interest rate differential relative to the US for the following countries: Australia, Canada, the Eurozone (Germany before its introduction), New Zealand, Japan, Switzerland, and the United Kingdom. The weights are constructed as the value of each country's exports and imports as a fraction of the average value of trade over the seven countries. Monthly exchange rate data is from Datastream and interest rate differentials are calculated using covered interest rate parity (CIP) from forward rates, $i_t^\star - i_t = f_t - y_t$, also available from Datastream. We then run the following family of regressions:

$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^\star - i_t) + \varepsilon_{t,h},$$

where h is the horizon measured in months. This regression is, of course, identical to the return predictability regression (6) in Section 2.

Figure 1 plots the estimated slope coefficients at various horizons. For small values of h, we find the slope coefficient to be positive, thus recovering the classic forward discount puzzle: at short time horizons, higher interest rate differentials (relative to the US) lead to higher excess returns. This pattern remains the same up to a horizon of 3 years but then reverses sign, illustrating the predictability reversal puzzle: for horizons between 4 and 7 years, higher interest rate differentials predict lower excess returns. Finally, as the figure indicates, the estimated coefficient β_h^{rx} becomes indistinguishable from zero at even longer horizons.

5.2. *Return predictability*

Turning to our framework's predictions, we first calculate the ACF of the interest rate differential between the trade-weighted average interest rate and the US interest rate. Taking this autocorrelation, $\{\xi_\tau^*\}_{\tau\geq 1}$, as our primitive, we then calculate the term structure of the model-implied coefficients β_h^{rx} of the return predictability regression using expression (13) in Theorem 1 for different number of factors, k, in the agents' model.

As our first exercise, we consider the case in which agents can only entertain single-factor models, that is, $k = 1$. Recall that in this special case, we can use the closed-form expression (15) in Proposition 6 to calculate the model-implied slope coefficients of the return predictability regression. Figure 2 plots the term structure of the model-implied coefficients for $k = 1$ together with the coefficients obtained from the data from Figure 1. As the figure indicates, the pattern of model-implied coefficients tracks the pattern observed in the data fairly closely (though not

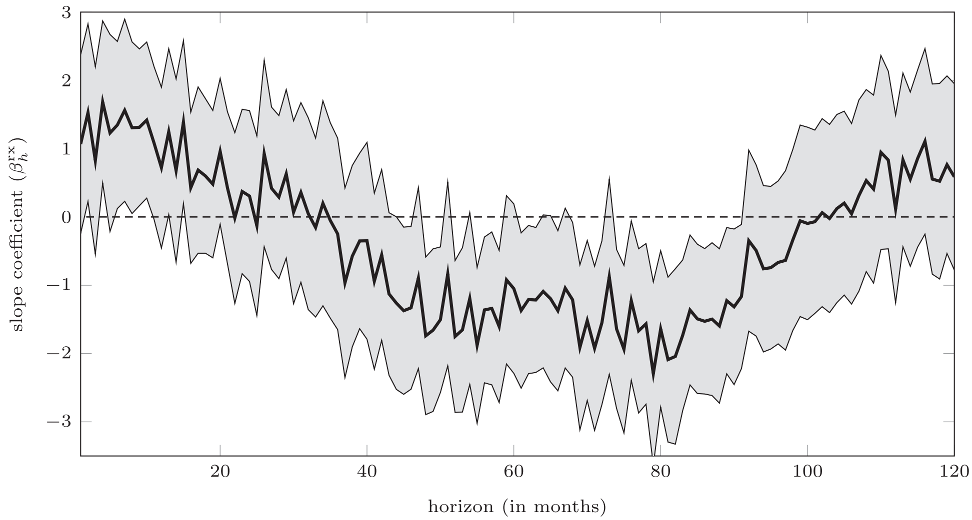

FIGURE 1
Violation of UIP at different Horizons

Notes: This figure plots estimated slope coefficients, together with 90% confidence intervals, of return predictability regression: $\text{rx}_{t+h} = \alpha_h^{\text{rx}} + \beta_h^{\text{rx}}(i_t^\star - i_t) + \varepsilon_{t,h}$, where h denotes the horizon in months, rx_{t+h} is the currency excess return on a trade-weighted average of seven currencies vis-à-vis the US dollar, and $i_t^\star - i_t$ is the log interest rate differential between the trade-weighted average interest rate and the US interest rate. Monthly data from January 1985 to December 2021. Confidence intervals are calculated using Newey and West (1987) standard errors.

its magnitude).[15] Most importantly, we see a reversal in the slope coefficient: model-implied coefficients are positive for horizons up to 28 months and reverse to a negative sign thereafter. Figure 2 thus suggests that if agents are constrained to rely on the family of single-factor models to make forecasts about the evolution of the interest rate differential process, their subjective expectations result in return predictability patterns that are simultaneously consistent with the forward discount and predictability reversal puzzles.

It is important to emphasize that while the estimated coefficients in Figure 1 are obtained from regressing returns on interest rate differentials, the model-implied coefficients in Figure 2 do not use the data on exchange rates or excess returns. Rather, they are simply obtained by plugging the autocorrelation of the interest rate differential into equation (15).

Next, we investigate how increasing the number of factors in agents' models impacts the term structure of model-implied slope coefficients. To this end, we once again use the empirical ACF of the interest rate differential as an input to calculate β_h^{rx} for different values of k. However, when $k > 1$, there is no closed-form expression for the model-implied slope coefficients. As a result, we use the characterization result in Theorem 1 to solve for β_h^{rx} numerically. As a by-product, we also obtain the model-implied ACF from the perspective of agents who are constrained to the class of k-factor models (equation (11)).[16]

The results are reported in Figure 3. The left panel depicts the model-implied autocorrelation for $k = 1, 2, 3$, and 10, together with the empirical autocorrelation of the trade-weighted average

15. The mismatch between the magnitudes is due to the fact that the misspecification in agents' beliefs is the only driver of currency excess returns in the model, as the reduced-form formulation in equation (2) abstracts from (subjective) risk premia.

16. Note that the expressions in equations (11)–(13) involve infinite sums. To implement these expressions numerically, we truncate the infinite sums after 120 terms.

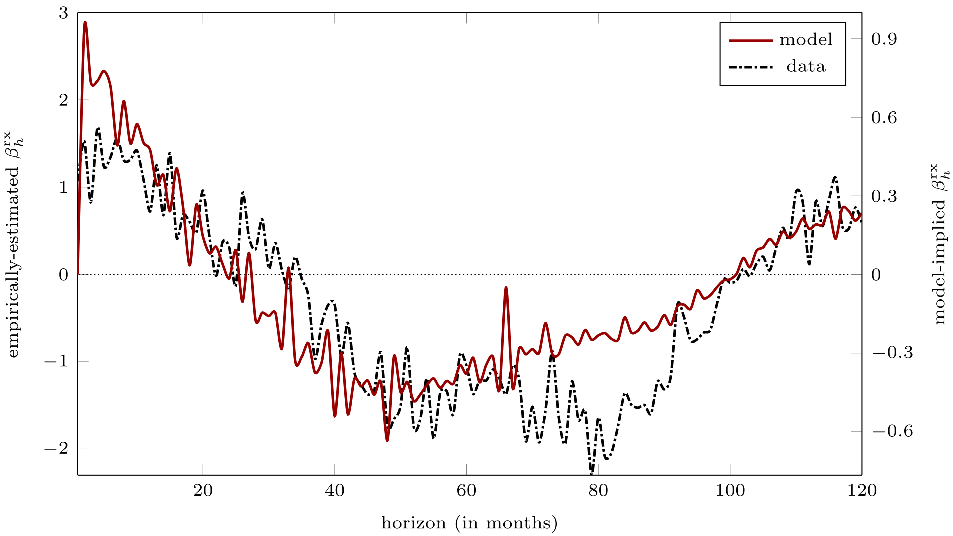

FIGURE 2
Term structure of return predictability

Notes: This figure plots estimated slope coefficients of return predictability regression from Figure 1 (left axis) together with model-implied coefficients from a one-factor model given in Proposition 6 (right axis) for different horizons. Monthly data from January 1985 to December 2021.

interest rate differential. As the figure indicates, the model-implied ACFs can differ quite substantially across various levels of agents' sophistication. For example, while the 10-factor model exhibits patterns that are fairly similar to the empirical ACF, the ACF implied by the single-factor model looks significantly different. Crucially, this is reflected in the model-implied slope coefficients as illustrated in the right panel of Figure 3: the 10-factor model, which generates a model-implied ACF that closely tracks the empirical ACF, also results in model-implied slope coefficients that are significantly smaller at nearly all horizons. This, of course, is to be expected in view of our results in Section 4. As agents are able to entertain richer and more complex statistical models, they end up with models that better fit the empirical ACF, which in turn results in less significant deviations from the rational expectations benchmark and hence less return predictability.

5.3. *Forecast-error predictability*

The findings in Subsection 5.2 indicate that our behavioural framework can generate return predictability patterns that are simultaneously consistent with the forward discount and predictability reversal puzzles. In this subsection, we investigate whether the predictions of our behavioural framework are also consistent with evidence from survey data. Specifically, we focus on the dynamics of forecast errors and compare the coefficients of forecast-error predictability regression (7) implied by our model to their empirical counterparts from survey data.

To obtain the empirical counterpart of the left-hand side of equation (7), we rely on surveys of professional forecasters, provided to us by *Consensus Economics*. The surveys query respondents every month about a number of country-specific macroeconomic and financial variables, including 3-month-ahead forecasts of 3-month interest rates in different countries (Consensus Economics Inc., 2022). Using these surveys, we calculate a "consensus forecast" for 3-month

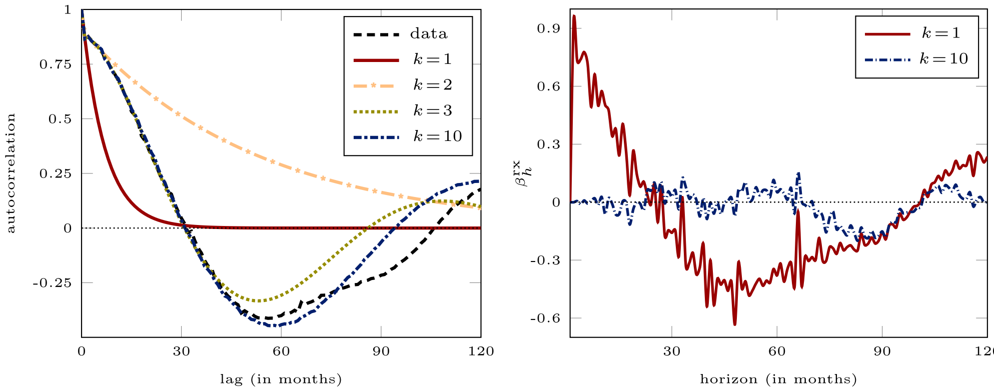

FIGURE 3

Varying the complexity of agents' model

Notes: The left panel plots the ACF of the trade-weighted interest rate differential (data) together with model-implied autocorrelations for models with $k = 1, 2, 3$, and 10 factors constructed from equation (11). The right panel plots the model-implied slope coefficients of the return predictability regression (6) at different horizons for models with $k = 1$ and $k = 10$ factors. Monthly data from January 1985 to December 2021.

interest rates for Canada, the Eurozone, Japan, Switzerland, the United Kingdom, and the United States by taking the median of all forecasts for each country.[17,18] We then use these consensus forecasts to construct forecasts for interest rate differentials between the United States and each of the other five countries at a monthly frequency. Finally, we subtract the resulting forecasts from the realized interest rate differentials and obtain monthly estimates for forecast errors of interest rate differentials.

To estimate the slope coefficients of the family of forecast-error predictability regressions, we regress the trade-weighted forecast errors of interest rate differentials on the trade-weighted realized interest rate differentials. To calculate the model-implied slope coefficients, we consider the case that agents rely on a single-factor model (*i.e.* $k = 1$) and use the expression in equation (16) for $m = 3$. Figure 4 plots the empirical and model-implied $\beta^{\mathrm{fe}}_{h,3}$ at various horizons. As is evident from the figure, under the assumption that $k = 1$, the model-implied slope coefficients of forecast-error predictability regression (7) track their empirical counterparts fairly closely. Note that whereas the empirically estimated coefficients in Figure 4 are constructed from survey data, the model-implied coefficients are obtained by simply plugging the autocorrelation of the interest rate differential into equation (16).

As a final exercise, we use our result in Proposition 8 to obtain a direct estimate for the number of factors that would be consistent with survey data. Recall from equation (18) that, as long as $k \leq n$, the rank of the Hankel matrix $\mathbf{\Phi}$ constructed from the coefficients in (17) is simply equal to k. We can therefore perform a direct test of the value of k using the following three-step procedure: (i) regress the consensus (3-month-ahead) forecasts of interest rate differential obtained

17. Our data set does not contain forecasts for Australia and New Zealand. We therefore drop these countries from the rest of the analysis. We also restrict our attention to monthly data from June 1998 to December 2021, which is the period over which we have access to a cross-section of different forecasts for each of the countries in our sample. See Supplementary Appendix E for more details about the data set.

18. Also see Stavrakeva and Tang (2020) and Kalemli-Özcan and Varela (2023) for other recent studies of determinants of exchange rate dynamics using survey data and Piazzesi *et al.* (2015) who use survey data to study the importance of investor beliefs for bond return predictability.

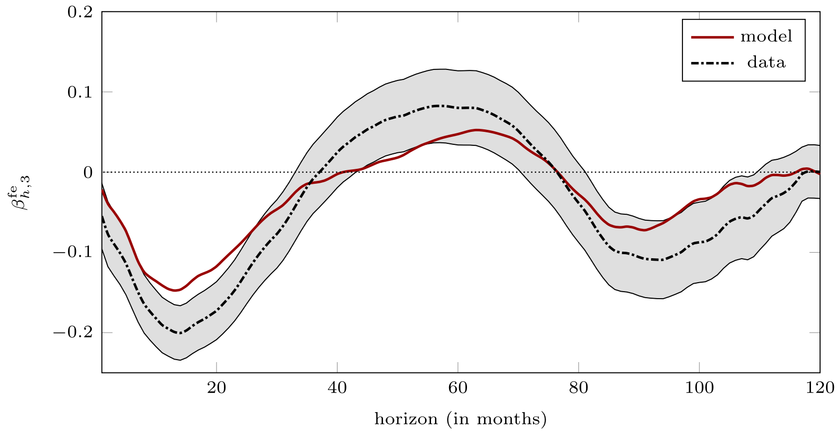

FIGURE 4

Term structure of forecast-error predictability.

Notes: This figure plots the empirically estimated (together with 95% confidence intervals) and model-implied slope coefficients of forecast-error predictability regression (7) for forecast horizon $m = 3$ months. The model-implied coefficients are constructed under the assumption that $k = 1$ from equation (16). Monthly data from June 1998 to December 2021.

from survey data on the past realizations of the (1-month CIP-implied) interest rate differential to estimate coefficients $(\phi^{(3)}_1, \phi^{(3)}_2, \ldots)$ in equation (17); (ii) construct the Hankel matrix $\mathbf{\Phi}$ with typical element $\mathbf{\Phi}_{ij} = \phi^{(3)}_{i+j-1}$; and (iii) test for the rank of $\mathbf{\Phi}$. Note that, technically speaking, implementing step (i) requires regressing the forecasts of interest rate differentials on all past realizations of interest rate differentials. To implement this step empirically, we perform the above procedure while varying the number of lags used for estimating (17) from 10 to 20.

Irrespective of the number of lags used, we cannot reject the null hypothesis that $\text{rank}(\mathbf{\Phi}) = 1$: the p-values corresponding to the rank test of Donald *et al.* (2007) never fall below 78%.[19] In view of Proposition 8, this finding indicates that forecasts of interest rate differentials are consistent with a model in which forecasts are formed using a single-factor model. This is, of course, also consistent with the findings reported in Figure 4 for the extent of forecast-error predictability at different horizons.

6. CONCLUSIONS

This paper studies how limits to the complexity of statistical models that agents use for forecasting shape asset prices, where we define the complexity of a statistical model as the dimension of its minimal representation. We develop our results in the context of a simple framework in which a sequence of exogenous fundamentals are generated by a stochastic process that may be more complex than what agents can entertain. As a result, agents form their subjective expectations by relying on lower-dimensional approximations to the underlying data-generating process. The constraint on the complexity of agents' models is our only point of departure from the textbook rational expectations framework: we impose no-arbitrage, maintain the assumption of Bayesian updating, and assume that agents exhibit no other behavioural biases.

19. See Supplementary Appendix F for details. We also perform the same exercise for each country separately, finding similar results.

As our main theoretical result, we characterize how the statistical properties of the true data-generating process together with the limit on the number of factors in agents' models shape agents' forecasts and the resulting return dynamics. Our theoretical results provide us with sharp characterizations of the slope coefficients of the return and forecast-error predictability regressions at all horizons. We then apply our framework to study violations of UIP in foreign exchange markets. We find that the deviations from rational expectations implied by the limit on the complexity of agents' statistical models can generate the well-known forward discount and predictability reversal puzzles, while simultaneously being consistent with forecast-error predictability patterns from survey data.

In order to obtain a tractable framework, we made a number of simplifying assumptions. First, we focused on how relying on lower-dimensional approximations to the data-generating process shapes the term structure of return predictability for a single asset. However, phenomena such as momentum and reversal are pervasive, not just in the time series, but also in the cross section. Extending our framework to multiple assets and exploring the possible implications of constraints on model complexity for cross-sectional return predictability would be a natural next step for future work. Second, we focused on a simple environment in which prices relate to fundamentals via the simple no-arbitrage condition in equation (2). While this allowed us to isolate the implications of agents' inability to entertain high-dimensional models for return dynamics in a transparent manner, we abstracted from other important features, such as time-varying (fundamental) risk premia. Integrating our behavioural assumption into a more structural setting with general preferences would shed further light on how distortions in subjective expectations caused by constraints on model complexity impact asset prices. Finally, throughout the paper, we assumed k to be exogenous, thus treating the maximum dimension of agents' model as a behavioural primitive. An alternative interpretation of the complexity constraint is technological: agents can in principle use models of any dimension but find estimating a higher-dimensional model to be more costly. Exploring the implications of this alternative interpretation requires extending our framework, whereby market participants endogenously choose the complexity of their models while trading off the benefit of a more accurate understanding of the data-generating process with the cost of estimating a higher-dimensional model.

APPENDIX

A. Additional results

A.1. *Model complexity and bounded recall*

The behavioural assumption in the paper maintains that agents can only hold and update beliefs over the set of models with at most k hidden factors, as represented by equation (4). In this subsection, we show that an agent who is restricted to the class of k-factor models is isomorphic to an agent with bounded recall, who relies on at most k moving averages as summary statistics for the entire history of her past observations to make predictions about the future. We start by formalizing how such an agent uses her past observations to make forecasts. We then establish the isomorphism between the agent with limited memory and an agent who is constrained to using k-factor models.

Consider an agent who, at any given time t, constructs k moving averages $s_t = (s_{1t}, s_{2t}, \dots, s_{kt})$ of the past realizations of the fundamental:

$$s_{it} = w_i x_t + \sum_{j=1}^{k} q_{ij} s_{jt-1}, \tag{A.1}$$

where x_t is the realization of the fundamental at time t, and $w = (w_1, \ldots, w_k)$ and $\mathbf{Q} = [q_{ij}] \in \mathbb{R}^{k\times k}$ are a vector and a matrix of constants, respectively. We assume that one of these statistics is simply equal to the fundamental's most recent realization.[20] The key assumption is that the agent treats these moving averages as summary statistics for making predictions about all future realizations of the fundamental:

$$\mathbb{E}_t[x_{t+\tau}] = \sum_{i=1}^{k} v_{\tau i} s_{it} \quad \text{for all } \tau \geq 1, \tag{A.2}$$

where $\mathbb{E}_t[\cdot]$ denotes the agent's subjective expectation and the sequence of vectors $v_\tau = (v_{\tau 1}, \ldots, v_{\tau k})$ determines how the agent maps the statistics she maintains about the past into forecasts about the future. Throughout, we assume that the agent's subjective expectations in equation (A.2) are internally consistent and satisfy the law of iterated expectations.

A few remarks are in order. First, note that equations (A.1) and (A.2) represent an agent who is restricted to forecasting the entire future path of the fundamental by relying on at most k statistics of her past observations. Such an agent can thus be interpreted as one with bounded recall: an agent with a larger k can construct a richer set of statistics and hence maintain a more complete record of the past, with the case of $k \to \infty$ corresponding to an agent who can, in principle, recall the entire history of her observations.[21] Second, it is immediate to verify that the formulation in equation (A.1) nests the framework in Fuster *et al.* (2010, 2012), who consider an agent who can only recall her k most recent observations. In contrast to their set-up, the agent represented by equation (A.1) can also construct richer statistics that are conditioned on long moving averages of observations, rather than only the most recent history. Finally, note that, aside from internal consistency of subjective expectations, the formulation in equation (A.2) does not impose any restrictions on how the agent uses her moving averages to form forecasts about the future; all that matters is that forecasts for all horizons are constructed using the same k statistics maintained by the agent.

To state our next result, recall that the set of hidden k-factor models studied in the main body of the paper can be parameterized by $\theta \in \Theta_k$, where $\theta = (\mathbf{A}, \mathbf{B}, c)$ denotes the model parameters in equation (4). We similarly parameterize the set of agents who forecast the future using k moving averages by $\psi \in \Psi_k$, where $\psi = (w, \mathbf{Q}, \{v_\tau\}_{\tau=1}^{\infty})$ denotes the collection of parameters in equations (A.1) and (A.2). Additionally, we say two agents are *isomorphic* if, given the same (infinitely long) history of observations, their forecasts for future realizations of the fundamental coincide with one another at all horizons. We have the following result.

Proposition A.1. *There is an isomorphism between Θ_k and Ψ_k.*

The above result establishes that, given an agent who uses a k-factor model with parameters $(\mathbf{A}, \mathbf{B}, c) \in \Theta_k$, there always exists an agent with parameters $(w, \mathbf{Q}, \{v_\tau\}_{\tau=1}^{\infty}) \in \Psi_k$ who, when faced the same history of observations, makes the same forecasts at all horizons by maintaining k moving averages of her past observations (and vice versa).

20. Specifically, we assume that $s_{1t} = x_t$ for all t. This means that $w_1 = 1$ and $q_{1j} = 0$ for all j. We also assume that w and $\mathbf{Q}$ are such that (i) the vector of moving averages s_t remains bounded at all times and (ii) each element of s_t contains some information about the fundamental (*i.e.* there does not exist an i such that $s_{it} = 0$ for all t).

21. Note that the notion of bounded recall represented by equations (A.1) and (A.2) is different from the similarity-based recall model of Bordalo *et al.* (2020) and the imprecise memory model of da Silveira *et al.* (2020). Unlike these models, the behavioural assumption represented by equations (A.1) and (A.2) is that the agent can only rely on up to k moving averages for making forecasts, even when the true data-generating process is more complex and cannot be summarized via k statistics.

Proposition A.1 provides an alternative interpretation for the wedge between n and k and the restriction on the complexity of agents' models. In the main body of the paper, we interpreted an agent with $k < n$ as an agent who predicts the future by relying on a less complex statistical model than the true data-generating process. In view of Proposition A.1, an agent with $k < n$ can alternatively be interpreted as an agent who summarizes the past via fewer statistics than what is necessary for making unbiased forecasts about the future. Put differently, the restriction on the complexity of an agent's model of the data-generating process is equivalent to imposing a restriction on the number statistics the agent can maintain to recall the past.

A.2. *Non-Bayesian estimation*

Throughout the paper, we maintained that agents form their subjective expectations by learning from the past realizations of the fundamental and updating their beliefs in a Bayesian fashion. Additionally, in Proposition 2, we showed that Bayesian learning implies that agents' posterior beliefs concentrate on the subset of k-factor models with minimum KL divergence from the true data-generating process. In this subsection, we show that agents end up with the same set of models if, instead of Bayesian updating, they choose their model to minimize the mean-squared error of their forecasts or rely on a maximum likelihood estimator.

Proposition A.2. *Consider an agent who is restricted to the class of k-factor models, Θ_k, and let $\widehat{\Theta}_k$ denote an arbitrary compact subset of Θ_k.*

(1) Let $\hat{\theta}_t^{\mathrm{ML}} \in \arg\max_{\theta\in\widehat{\Theta}_k} f^\theta(x_t,\ldots,x_0)$ denote the agent's maximum likelihood estimator. Then,

$$\lim_{t\to\infty} \mathrm{KL}(\theta^*\|\hat{\theta}_t^{\mathrm{ML}}) = \min_{\theta\in\widehat{\Theta}_k} \mathrm{KL}(\theta^*\|\theta) \quad \mathbb{P}^*\text{-almost surely}.$$

(2) Let $\hat{\theta}_t^{\mathrm{MSE}} \in \arg\min_{\theta\in\widehat{\Theta}_k} \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^\theta[x_{\tau+1}])^2$ denote a model that minimizes the time-series average of the agent's squared forecast errors. Then,

$$\lim_{t\to\infty} \mathrm{KL}(\theta^*\|\hat{\theta}_t^{\mathrm{MSE}}) = \min_{\theta\in\widehat{\Theta}_k} \mathrm{KL}(\theta^*\|\theta) \quad \mathbb{P}^*\text{-almost surely}.$$

This result establishes that, under either maximum likelihood or mean-squared error estimation, agents' estimates of the parameters of the data-generating process converge to the set of models with minimum KL divergence from the true data-generating process. Therefore, as long as agents have access to a large enough sequence of observations, they make forecasts using the same k-factor model irrespective of whether they rely on Bayesian updating, mean-squared-error minimization, or maximum likelihood estimation. This means that our results on the extent of return predictability in Theorem 1 and Propositions 3–8 remain applicable irrespective of which method agents use for model selection.[22]

A.3. *Heterogenous-agent economy*

Our results in the main body of the paper rely on the assumption that the economy consists of a unit mass of identical agents, all of whom are restricted to using models with the same maximum

22. The equivalence between Bayesian updating, maximum likelihood estimation, and mean-squared error minimization only holds when agents have access to an asymptotically large sample of observations. Given any finite sample, a Bayesian agent holds a non-degenerate posterior belief over the set of all k-factor models. The same clearly does not apply to the frequentist agents of Proposition A.2.

number of factors, k. In this subsection, we extend our results by assuming that only a fraction $1-\gamma$ of agents are subject to our behavioural constraint, while the remaining γ fraction can entertain models with any number of factors. For simplicity, we refer to the two groups of agents as behavioural and rational agents, respectively.[23]

The heterogeneity in the agents' ability to entertain statistical models of different complexities results in heterogenous subjective expectations. Therefore, as in Allen *et al.* (2006), we assume that the relationship between endogenous prices and exogenous fundamentals is given by the following generalization of equation (2):

$$y_t = x_t + \delta \bar{\mathbb{E}}_t[y_{t+1}], \tag{A.3}$$

where $\bar{\mathbb{E}}[\cdot] = \gamma\,\mathbb{E}^*[\cdot] + (1-\gamma)\mathbb{E}[\cdot]$ denotes the cross-sectional average of agents' expectations.[24] Iterating on the above, we can also obtain the following counterpart to equation (5) for excess returns in terms of agents' expectations:

$$\mathrm{rx}_{t+1} = \sum_{\tau=1}^{\infty} \delta^{\tau} \left(\bar{\mathbb{E}}_{t+1}\bar{\mathbb{E}}_{t+2}\cdots\bar{\mathbb{E}}_{t+\tau}[x_{t+\tau}] - \bar{\mathbb{E}}_{t}\bar{\mathbb{E}}_{t+1}\cdots\bar{\mathbb{E}}_{t+\tau}[x_{t+\tau}] \right). \tag{A.4}$$

The key observation is that, even though subjective expectations of each group of agents satisfy the law of iterated expectations, the cross-sectional average expectation $\bar{\mathbb{E}}[\cdot]$ may not. Therefore, unlike the representative-agent framework of Section 2, excess returns in the heterogenous-agent economy also depend on higher-order expectations whenever $k<n$. The failure of the law of iterated expectations with respect to $\bar{\mathbb{E}}[\cdot]$ in our framework resembles a similar phenomenon in differential-information economies, such as in Angeletos and Lian (2018) and Angeletos and Huo (2021).

To characterize the extent of return predictability in the heterogenous-agent economy, we assume that while rational agents can recover the model used by behavioural agents, behavioural agents behave as if they live in a representative-agent economy only consisting of agents with k-factor models. This assumption captures the idea that, given their priors, behavioural agents are convinced—mistakenly so when $k<n$—that a k-factor model is sufficient to capture the process that drives the fundamental.

Proposition A.3. *Let $\beta_h^{\mathrm{rx}}(\gamma)$ denote the slope coefficient of the return predictability regression (6) in an economy with γ and $1-\gamma$ fraction of rational and behavioural agents, respectively. Then,*

$$\beta_h^{\mathrm{rx}}(\gamma) = (1-\gamma)\sum_{s=0}^{\infty}(\delta\gamma)^s \beta_{h+s}^{\mathrm{rx}}(0), \tag{A.5}$$

where $\beta_h^{\mathrm{rx}}(0)$ is the slope coefficient in the representative-agent economy, given by equation (13).

23. Despite the terminology, recall that all agents in this economy are Bayesian, with the only difference between the two groups being that the "behavioural" agents assign zero prior beliefs to models consisting of more than k factors. Furthermore, note that when $k \geq n$, agents in both groups end up with identical subjective expectations. The subjective expectations of the two groups differ only when $k<n$.

24. We assume that $\bar{\mathbb{E}}[\cdot]$ in equation (A.3) is the unweighted cross-sectional average of agents' expectations. See Supplementary Appendix D for a simple micro-foundation based on the overlapping generations model of Allen *et al.* (2006). More generally, depending on the underlying micro-founded model, $\bar{\mathbb{E}}[\cdot]$ may be a weighted average of the agents' subjective expectations, with endogenous (and potentially state-dependent) weights (Panageas, 2020). Given our focus on a reduced-form framework, we abstract from these issues.

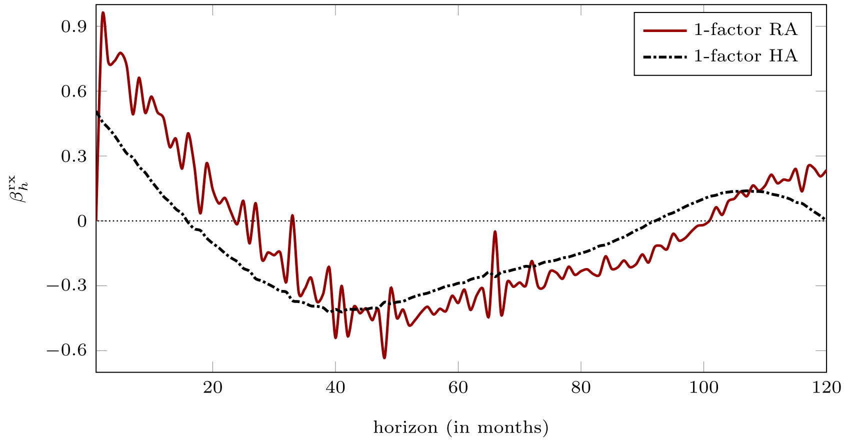

FIGURE A.1

Term structures of return predictability in the heterogeneous-agent economy

Notes: This figure plots the model-implied slope coefficients of the return predictability regression (6) at different horizons for the representative- and heterogenous-agent economies. Behavioural agents in both economies use a single-factor model. The fraction of behavioural agents in the heterogenous-agent economy is 10%, that is, $\gamma = 0.9$.

The above result relates the term structure of β_h^{rx} in the heterogenous-agent economy to that of the economy only consisting of behavioural agents. This allows us to use our results in the main body of the paper to characterize the extent and nature of return predictability in economies consisting of both rational and behavioural agents. Note that, according to equation (A.5), $\beta_h^{\mathrm{rx}}(\gamma)$ is not simply a weighted average of $\beta_h^{\mathrm{rx}}(0)$ and $\beta_h^{\mathrm{rx}}(1) = 0$. Rather, the fact that cross-sectional average expectations $\bar{\mathbb{E}}[\cdot]$ do not satisfy the law of iterated expectations implies that the extent of return predictability at horizon h in the heterogenous-agent economy also depends on the extent of return predictability in the representative-agent economy at all horizons $\tau \geq h$.

We conclude by studying the implications of Proposition A.3 in the context of the foreign exchange application in Section 5. Specifically, we use the expression in equation (A.5) to calculate the model-implied slope coefficients of the return predictability regression in a heterogeneous-agent economy, in which fraction $1-\gamma$ of agents are constrained to using a single-factor model, while the remaining γ fraction can entertain models with any number of factors. Figure A.1 plots the model-implied coefficients in an economy populated by 90% rational and 10% behavioural agents, that is, for $\gamma = 0.9$. As the figure illustrates, the slope coefficients in the heterogeneous-agent economy look similar to those in the representative-agent economy consisting of only behavioural agents (*i.e.* $\gamma = 0$). This indicates that even small fractions of behavioural agents can lead to notable deviations from the rational expectations benchmark.

B. Technical appendix

Proof of Proposition 1

According to agents' subjective model, the fundamental is generated by a stationary process with a representation in the form of equation (4). Therefore, we can express agents' m-step-ahead forecast of the fundamental as follows:

$$\mathbb{E}_{t-1}[x_{t+m-1}] = \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} x_{t-\tau}, \tag{B.1}$$

where $(\phi_1^{(m)}, \phi_2^{(m)}, \ldots)$ is a sequence of coefficients that depends on the parameters of the agents' model, $\theta = (\mathbf{A}, \mathbf{B}, c)$. Multiplying both sides of the above equation by x_{t-r}, taking unconditional expectations $\mathbb{E}[\cdot]$, and dividing both sides by $\mathbb{E}[x_t^2]$ implies that

$$\xi_{r+m-1} = \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} \xi_{r-\tau} \tag{B.2}$$

for all $r \geq 1$, where $\xi_\tau = \mathbb{E}[x_t x_{t+\tau}]/\mathbb{E}[x_t^2]$ is the autocorrelation of the fundamental at lag τ as perceived by the agents given their subjective model. Equation (B.2) thus relates the coefficients in equation (B.1) to the subjective autocorrelations implied by agents' model.

With equations (B.1) and (B.2) in hand, we now derive the expressions for the slope coefficients of the predictability regressions (6) and (7). Starting with the forecast-error predictability regression (7), the representation in equation (B.1) implies that

$$\beta_{h,m}^{\text{fe}} = \frac{1}{\mathbb{E}^*[x_t^2]} \left(\mathbb{E}^*[x_t x_{t+h+m-1}] - \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} \mathbb{E}^*[x_t x_{t+h-\tau}] \right) \tag{B.3}$$

for all $h \geq 1$, where $\mathbb{E}^*[\cdot]$ denotes the expectation under the true data-generating process. As a result,

$$\beta_{h,m}^{\text{fe}} = (\xi_{h+m-1}^* - \xi_{h+m-1}) - \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} (\xi_{h-\tau}^* - \xi_{h-\tau}) + \left(\xi_{h+m-1} - \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} \xi_{h-\tau} \right),$$

where $\xi_\tau^* = \mathbb{E}^*[x_t x_{t+\tau}]/\mathbb{E}^*[x_t^2]$ is the autocorrelation of the fundamental at lag τ. Note that, according to equation (B.2), the last term on the right-hand side of the above equation is equal to zero for all $h \geq 1$. As a result,

$$\beta_{h,m}^{\text{fe}} = (\xi_{h+m-1}^* - \xi_{h+m-1}) - \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} (\xi_{h-\tau}^* - \xi_{h-\tau}). \tag{B.4}$$

Note that we can write (B.2) as $\xi^{(m)} = \Xi \phi^{(m)}$, where $\xi^{(m)}$ and $\phi^{(m)}$ are infinite-dimensional vectors with elements $\xi_i^{(m)} = \xi_{m+i-1}$ and $\phi_i^{(m)}$, respectively, and Ξ is an infinite-dimensional matrix such that $\Xi_{ij} = \xi_{i-j}$ for all $i, j \geq 1$. Consequently, $\phi_\tau^{(m)} = \sum_{s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_{s+m-1}$. Replacing for $\phi_\tau^{(m)}$ into equation (B.4) then establishes equation (9).

Turning to the slope coefficient of the return predictability regression (6), note that by the law of iterated expectations, $\mathbb{E}_t[x_{t+\tau}] = \mathbb{E}_t \mathbb{E}_{t+\tau-1}[x_{t+\tau}] = \sum_{r=1}^{\infty} \phi_r \mathbb{E}_t[x_{t+\tau-r}]$ for all $\tau \geq 1$, where we are using the representation in equation (B.1) and for notational simplicity denoting $\phi_r^{(1)}$ by ϕ_r. Consequently,

$$\begin{aligned} \mathbb{E}_{t+1}[x_{t+\tau}] - \mathbb{E}_t[x_{t+\tau}] &= \sum_{r=1}^{\infty} \phi_r (\mathbb{E}_{t+1}[x_{t+\tau-r}] - \mathbb{E}_t[x_{t+\tau-r}]) \\ &= \sum_{r=1}^{\tau-1} \phi_r (\mathbb{E}_{t+1}[x_{t+\tau-r}] - \mathbb{E}_t[x_{t+\tau-r}]) \end{aligned}$$

for all $\tau \geq 2$. Multiplying both sides of the above equation by δ^τ, summing over all $\tau \geq 2$, and using the expression for excess returns in equation (5) implies that

$$\begin{aligned}\mathrm{rx}_{t+1} &= \delta\left(x_{t+1} - \mathbb{E}_t[x_{t+1}]\right) + \sum_{\tau=2}^{\infty} \delta^\tau \sum_{r=1}^{\tau-1} \phi_r \left(\mathbb{E}_{t+1}[x_{t+\tau-r}] - \mathbb{E}_t[x_{t+\tau-r}]\right) \\ &= \delta\left(x_{t+1} - \mathbb{E}_t[x_{t+1}]\right) + \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \delta^{r+s} \phi_r \left(\mathbb{E}_{t+1}[x_{t+s}] - \mathbb{E}_t[x_{t+s}]\right).\end{aligned}$$

Using equation (5) one more time, we can express the second term on the right-hand side of the above equation in terms of excess returns. In particular, $\mathrm{rx}_{t+1} = \delta(x_{t+1} - \mathbb{E}_t[x_{t+1}]) + \mathrm{rx}_{t+1} \sum_{r=1}^{\infty} \delta^r \phi_r$, which in turn implies that $\mathrm{rx}_{t+1} = \delta(x_{t+1} - \mathbb{E}_t[x_{t+1}])/(1 - \sum_{r=1}^{\infty} \delta^r \phi_r)$. As a result, we can express the slope coefficients of the return predictability regression in terms of the slope coefficients of the forecast-error predictability regression as follows:

$$\beta_h^{\mathrm{rx}} = \frac{\delta \beta_{h,1}^{\mathrm{fe}}}{1 - \sum_{r=1}^{\infty} \delta^r \phi_r} \quad \text{for all } h \geq 1. \tag{B.5}$$

Replacing for $\beta_{h,1}^{\mathrm{fe}}$ from equation (9) and using the fact that $\phi_\tau = \sum_{s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_s$ establishes equation (8). □

Proof of Proposition 2

We prove this result by verifying that our setting satisfies Assumptions 1–7 of Shalizi (2009) and invoking Theorem 3 of that paper. To this end, let

$$\ell^\theta(x_t, \ldots, x_0) = \log \frac{f^*(x_t, \ldots, x_0)}{f^\theta(x_t, \ldots, x_0)}$$

denote the log of the ratio of likelihoods of the fundamental under the true data-generating process (f^*) and under the agents' subjective model (f^θ). This object is clearly measurable with respect to time t information, thus implying that Assumption 1 of Shalizi (2009) is trivially satisfied.

Next, we verify that Assumptions 2–4 are also satisfied. By the uniform law of large numbers (White, 1994, Theorem A.2.2), $\lim_{t\to\infty} \frac{1}{t} \log f^\theta(x_t, x_{t-1}, \ldots, x_0) = h(\theta^*, \theta)$ with $\mathbb{P}^*$-probability one uniformly in $\theta \in \widehat{\Theta}_k$, where $h(\theta^*, \theta) = \mathbb{E}^*[\log f^\theta(x_{t+1}|x_t, x_{t-1}, \ldots)]$. This implies that

$$\lim_{t\to\infty} \frac{1}{t} \ell^\theta(x_t, x_{t-1}, \ldots, x_0) = \mathrm{KL}(\theta^* || \theta) \quad \mathbb{P}^*\text{-almost surely} \tag{B.6}$$

uniformly in $\theta \in \widehat{\Theta}_k$, where $\mathrm{KL}(\theta^*||\theta) = h(\theta^*, \theta^*) - h(\theta^*, \theta)$ is the KL divergence defined in equation (10). Equation (B.6) readily implies Assumption 3. Furthermore, note that equation (B.6), together with the dominated convergence theorem, implies that

$$\lim_{t\to\infty} \frac{1}{t} \mathbb{E}^*[\ell^\theta(x_t, \ldots, x_0)] = \mathrm{KL}(\theta^*||\theta),$$

thus establishing Assumption 2. Finally, note that Assumption 4 follows from the fact that, in our setting, the KL divergence is guaranteed to be finite.

We now turn to verifying that Assumptions 5–7 of Shalizi (2009) are also satisfied. To this end, let $G_t = \widehat{\Theta}_k$ for all t. This readily establishes Assumption 7. It also implies that $\mu_0(G_t) = 1$ and $\text{ess inf}_{\theta \in G_t} \text{KL}(\theta^* \| \theta) = \text{ess inf}_{\theta \in \widehat{\Theta}_k} \text{KL}(\theta^* \| \theta)$. Furthermore, (B.6) guarantees that

$$\lim_{t \to \infty} \sup_{\theta \in G_t} \left| \frac{1}{t} \ell^\theta(x_t, x_{t-1}, \ldots, x_0) - \text{KL}(\theta^* \| \theta) \right| = 0 \quad \mathbb{P}^*\text{-almost surely.}$$

Thus, all three statements of Assumption 5 are satisfied. Finally, to verify Assumption 6, note that setting $G = \widehat{\Theta}_k$ in equation (18) of Shalizi (2009) implies that

$$\limsup_{t \to \infty} \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \ldots, x_0)}{f^*(x_t, \ldots, x_0)} \mathrm{d}\mu_0(\theta) \leq - \min_{\theta \in \widehat{\Theta}_k} \text{KL}(\theta^* \| \theta) \quad \mathbb{P}^*\text{-almost surely.}$$

Therefore, $\tau(\widehat{\Theta}_k, \delta) \leq t$ with $\mathbb{P}^*$-probability one for all $\delta > 0$ and t sufficiently large, where

$$\begin{aligned} \tau(G, \delta) = \max \Bigg\{ t : \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \ldots, x_0)}{f^*(x_t, \ldots, x_0)} \mathbf{1}_{\{\theta \in G\}} \mathrm{d}\mu_0(\theta) \\ > \delta + \limsup_{t \to \infty} \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \ldots, x_0)}{f^*(x_t, \ldots, x_0)} \mathbf{1}_{\{\theta \in G\}} \mathrm{d}\mu_0(\theta) \Bigg\}. \end{aligned}$$

Thus, setting $G_t = \widehat{\Theta}_k$ ensures that Assumption 6 is also satisfied.

Having verified that Assumptions 1–7 of Shalizi (2009) are satisfied, Theorem 3 of that paper implies the statement in Proposition 2. □

Proof of Theorem 1

Proof of Part 1. By Proposition 2, agents' long-run beliefs concentrate on the set of models with minimum KL divergence from the true model. We thus start by characterizing $\arg\min_{\theta \in \Theta_k} \text{KL}(\theta^* \| \theta)$. Note that instead of optimizing over Θ_k, we can optimize over $\underline{\Theta}_1 \cup \underline{\Theta}_2 \cup \cdots \cup \underline{\Theta}_k$, where $\underline{\Theta}_r$ is the set of models whose minimal realization consists of r factors. Therefore, in what follows, and without loss of generality, we assume that model $\theta = (\mathbf{A}, \mathbf{B}, c)$ is a minimal realization consisting of $r \leq k$ factors.

Under model $\theta \in \underline{\Theta}_r$, agents believe that the fundamental is described by the process in equation (4), where $\omega_t \in \mathbb{R}^r$ is the vector of r hidden factors. As a result, conditional on $\{x_{t-\tau}\}_{\tau=0}^\infty$, agents believe that ω_{t+1} is normally distributed with mean $\hat{\omega}_t = \mathbb{E}_t[\omega_{t+1}]$ and variance $\hat{\Sigma}$, where $\hat{\Sigma}$ is the unique positive-definite matrix that satisfies the algebraic Riccati equation

$$\hat{\Sigma} = \mathbf{A} \left(\hat{\Sigma} - \frac{1}{c' \hat{\Sigma} c} \hat{\Sigma} c c' \hat{\Sigma} \right) \mathbf{A}' + \mathbf{B}\mathbf{B}', \tag{B.7}$$

$\hat{\omega}_t$ satisfied the recursive equation $\hat{\omega}_t = (\mathbf{A} - gc')\hat{\omega}_{t-1} + g x_t$, and $g \in \mathbb{R}^r$ is the Kalman gain given by

$$g = \mathbf{A} \hat{\Sigma} c (c' \hat{\Sigma} c)^{-1}. \tag{B.8}$$

Conditional on $\{x_{t-\tau}\}_{\tau=0}^\infty$, agents believe that the fundamental x_{t+1} is normally distributed with mean $\mathbb{E}_t[x_{t+1}] = c' \hat{\omega}_t$ and variance $\hat{\sigma}_x^2 = c' \hat{\Sigma} c$. Furthermore, their s-step-ahead forecasts of the future realization of the fundamental is given by

$$\mathbb{E}_t[x_{t+s}] = c'\mathbf{A}^{s-1}\sum_{\tau=0}^{\infty}(\mathbf{A}-gc')^{\tau}gx_{t-\tau} \tag{B.9}$$

for all $s \geq 1$, where g is the Kalman gain in equation (B.8). The above expression implies that the KL divergence (10) of agents' model θ from the true data-generating process is given by

$$\begin{aligned}\mathrm{KL}(\theta^*\|\theta) = &-\frac{1}{2}\log(\hat{\sigma}_x^{-2}) + \frac{1}{2}\log(2\pi) + \frac{1}{2}\hat{\sigma}_x^{-2}\mathbb{E}^*[x_t^2] - \hat{\sigma}_x^{-2}\sum_{s=1}^{\infty}c'(\mathbf{A}-gc')^{s-1}g\mathbb{E}^*[x_tx_{t-s}]\\ &+\frac{1}{2}\hat{\sigma}_x^{-2}\sum_{s=1}^{\infty}\sum_{\tau=1}^{\infty}c'(\mathbf{A}-gc')^{s-1}gc'(\mathbf{A}-gc')^{\tau-1}g\mathbb{E}^*[x_tx_{t+\tau-s}]\\ &+\mathbb{E}^*[\log f^*(x_{t+1}|x_t,\ldots)].\end{aligned} \tag{B.10}$$

To minimize the above over $\theta \in \underline{\Theta}_r$, it is convenient to work with the following change of variables:

$$\mathbf{M} = \hat{\Sigma}^{-1/2}\mathbf{A}\hat{\Sigma}^{1/2} \quad \text{and} \quad u = \frac{\hat{\Sigma}^{1/2}c}{\sqrt{c'\hat{\Sigma}c}}. \tag{B.11}$$

Note that $u \in \mathbb{R}^r$ has unit length and $\mathbf{M} \in \mathbb{R}^{r\times r}$ is a stable matrix. In view of equation (B.11), agents' s-step-ahead forecasts in equation (B.9) can be written as

$$\mathbb{E}_t[x_{t+s}] = u'\mathbf{M}^{s-1}\sum_{\tau=0}^{\infty}[\mathbf{M}(\mathbf{I}-uu')]^{\tau}\mathbf{M}ux_{t-\tau} \tag{B.12}$$

for all $s \geq 1$. Similarly, substituting for $\mathbf{A}$, c, and g in terms of $\mathbf{M}$ and u in (B.10) implies that

$$\begin{aligned}\mathrm{KL}(\theta^*\|\theta) = &\;\mathbb{E}^*[\log f^*(x_{t+1}|x_t,\ldots)] - \frac{1}{2}\log(\hat{\sigma}_x^{-2}) + \frac{1}{2}\log(2\pi)\\ &+\frac{1}{2}\hat{\sigma}_x^{-2}\mathbb{E}^*[x_t^2] - \hat{\sigma}_x^{-2}\sum_{s=1}^{\infty}\mathbb{E}^*[x_tx_{t-s}]u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u\\ &+\frac{1}{2}\hat{\sigma}_x^{-2}\sum_{s=1}^{\infty}\sum_{\tau=1}^{\infty}u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}uu'[\mathbf{M}(\mathbf{I}-uu')]^{\tau-1}\mathbf{M}u\mathbb{E}^*[x_tx_{t+\tau-s}].\end{aligned} \tag{B.13}$$

As we show in Lemma C.1 in the Supplementary Appendix, minimizing the above with respect to $(\mathbf{M}, u, \hat{\sigma}_x^{-2})$ and subject to the constraint that $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}') < 1$ is equivalent to minimizing (B.10) with respect to $(\mathbf{A}, \mathbf{B}, c)$. Therefore, to obtain the KL minimizer, we minimize equation (B.13) with respect to $(\mathbf{M}, u, \hat{\sigma}_x^{-2})$ subject to $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}') < 1$. We first minimize equation (B.13) with respect to $\hat{\sigma}_x^{-2}$. Taking the corresponding first-order condition and plugging back the result into equation (B.13) implies that minimizing equation (B.13) over $(\mathbf{M}, u, \hat{\sigma}_x^{-2})$

is equivalent to minimizing

$$\mathrm{KL}(\theta^*\|\theta) = \mathbb{E}^*[\log f^*(x_{t+1}|x_t,\ldots)] + \log\sqrt{2\pi e} + \frac{1}{2}\log\mathbb{E}^*[x_t^2] + \frac{1}{2}\log\left(1 - 2\sum_{s=1}^{\infty}\phi_s\xi_s^* + \sum_{\tau=1}^{\infty}\phi_s\phi_\tau\xi_{\tau-s}^*\right) \tag{B.14}$$

over $(\mathbf{M}, u)$, where ξ_s^* denotes the true autocorrelation of the fundamental at lag s and $\phi_s = u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$. Since only the last term on the right-hand side of equation (B.14) depends on θ, it follows immediately that minimizing $\mathrm{KL}(\theta^*\|\theta)$ is equivalent to minimizing equation (12) over $\mathbf{M}$ and u.

We next derive the expression in equation (11) for the autocorrelation of the fundamental according to agents' subjective model. Equation (4) implies that $\mathbb{E}[x_t x_{t-h}] = c'\mathbb{E}[\omega_t\omega_{t-h}']c = c'\mathbf{A}^h\mathbb{E}[\omega_t\omega_t']c$ for all $h \geq 0$. Therefore, applying the change of variables in equation (B.11) implies that the subjective autocorrelation at lag $h \geq 0$ is given by

$$\xi_h = \frac{u'\mathbf{M}^h\mathbf{G}u}{u'\mathbf{G}u}, \tag{B.15}$$

where $\mathbf{G} = \hat{\Sigma}^{-1/2}\mathbb{E}[\omega_t\omega_t']\hat{\Sigma}^{-1/2}$. To express $\mathbf{G}u$ in terms of $\mathbf{M}$ and u, note that (4) implies that $\mathbb{E}[\omega_t\omega_t'] = \mathbf{A}\mathbb{E}[\omega_t\omega_t']\mathbf{A}' + \mathbf{BB}'$. Multiplying both sides of this equation from the left and the right by $\hat{\Sigma}^{-1/2}$ implies that $\mathbf{G} = \mathbf{MGM}' + \hat{\Sigma}^{-1/2}\mathbf{BB}'\hat{\Sigma}^{-1/2}$. Consequently,

$$\mathbf{G} = \sum_{s=0}^{\infty}\mathbf{M}^s(\hat{\Sigma}^{-1/2}\mathbf{BB}'\hat{\Sigma}^{-1/2})\mathbf{M}'^s. \tag{B.16}$$

Next, note that multiplying both sides of the the algebraic Riccati equation in equation (B.7) by $\hat{\Sigma}^{-1/2}$ from the left and the right implies that $\hat{\Sigma}^{-1/2}\mathbf{BB}'\hat{\Sigma}^{-1/2} = \mathbf{I} - \mathbf{MM}' + \mathbf{M}uu'\mathbf{M}'$, where we are using the expressions in equation (B.11). Plugging this into the expression for $\mathbf{G}$ in equation (B.16) and multiplying both sides by u from the right leads to $\mathbf{G}u = u + \sum_{s=1}^{\infty}\mathbf{M}^s uu'\mathbf{M}^s u$. This, together with equation (B.15) and the fact that $u'u = 1$ then establishes equation (11). □

Proof of Part 2. According to equation (B.3), $\beta_{h,m}^{\mathrm{fe}} = \xi_{h+m-1}^* - \sum_{\tau=1}^{\infty}\phi_\tau^{(m)}\xi_{h-\tau}^*$, where $\{\phi_\tau^{(m)}\}_{\tau=1}^{\infty}$ are the coefficients in equation (B.1). Comparing equation (B.1) to the expression for agents' forecasts in equation (B.12) implies that $\phi_\tau^{(m)} = u'\mathbf{M}^{m-1}[\mathbf{M}(\mathbf{I}-uu')]^{\tau-1}\mathbf{M}u$ for all $\tau \geq 1$, thus establishing equation (14). To establish equation (13), recall from the proof of Proposition 1 that the slope coefficients of the return and forecast-error predictability regressions are related to one another via equation (B.5). Replacing for $\beta_{h,1}^{\mathrm{fe}}$ from equation (14) then establishes equation (13). □

Proof of Proposition 3

Proof of Part 1. When $k \geq n$, the true model is in the set of models considered by the agents, that is, $\theta^* \in \Theta_n \subseteq \Theta_k$. As a result, agents' subjective expectations coincide with rational expectations, which guarantees that $\xi_\tau = \xi_\tau^*$ for all τ. Thus, equations (8) and (9) imply that $\beta_h^{\mathrm{rx}} = \beta_{h,m}^{\mathrm{fe}} = 0$ for all $h \geq 1$. □

Proof of Part 2. We first show that if $\beta_{h,m}^{\mathrm{fe}} = 0$ for all h and m, then $k \geq n$. Suppose to the contrary that $k < n$. Let $(\mathbf{M}, u)$ and $(\mathbf{M}^*, u^*)$ denote the minimizers of equation (12) over the set of k- and n-factor models, respectively, and define $\phi_s = u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$ and $\phi_s^* =$

$u^{*\prime}[\mathbf{M}^*(\mathbf{I}-u^*u^{*\prime})]^{s-1}\mathbf{M}^*u^*$. By part 1, the slope coefficients of the forecast-error predictability regression arising from the optimal n-factor model are equal to zero at all horizons. Therefore, equation (14) implies that

$$\xi_h^* = \sum_{\tau=1}^{\infty} \phi_\tau^* \xi_{h-\tau}^* \quad \text{for all } h \geq 1. \tag{B.17}$$

Furthermore, by assumption, $\beta_{h,1}^{\text{fe}} = 0$ for all h under the k-factor model. Therefore, equation (14) implies that

$$\xi_h^* = \sum_{\tau=1}^{\infty} \phi_\tau \xi_{h-\tau}^* \quad \text{for all } h \geq 1. \tag{B.18}$$

Multiplying both sides of the first equation by ϕ_h and the second by ϕ_h^* and summing over all h, we get

$$\sum_{h=1}^{\infty} \phi_h^* \xi_h^* = \sum_{h=1}^{\infty} \phi_h \xi_h^*. \tag{B.19}$$

Next, note equations (B.17) and (B.18) also imply that the objective function (12) evaluated at the optimal solution in the set of all n- and k-factor models is, respectively, equal to

$$H(\mathbf{M}^*, u^*) = 1 - \sum_{s=1}^{\infty} \phi_s^* \xi_s^* \quad \text{and} \quad H(\mathbf{M}, u) = 1 - \sum_{s=1}^{\infty} \phi_s \xi_s^*.$$

Thus, by equation (B.19), $H(\mathbf{M}^*, u^*) = H(\mathbf{M}, u)$, which implies that the k-factor model results in the same KL divergence from the data-generating process as does the n-factor model, whose KL divergence from the true process is equal to zero by assumption. Therefore, the data-generating process has a representation with $k < n$ factors, which contradicts the assumption that n is the number of factors in the minimal representation of the data-generating process. Hence, if $\beta_{h,m}^{\text{fe}} = 0$ for all h and m, then $k \geq n$. This, coupled with equation (B.5), also guarantees that, if $k < n$, then there exists $h \geq 1$ such that $\beta_h^{\text{rx}} \neq 0$. □

Proof of Proposition 4

We start with two observations. First, the assumption that $\rho(\mathbf{A}^*) < 1$ implies that $\lim_{s\to\infty} \xi_s^* = 0$. Second, $\|\mathbf{M}(\mathbf{I}-uu')\|_2 = \sqrt{\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')} < 1$ guarantees that coefficients $\phi_s^{(m)} = u'\mathbf{M}^{m-1}[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$ in equations (13) and (14) are absolutely summable for any given m. In particular, the Cauchy–Schwarz inequality and the fact that the spectral norm is sub-multiplicative imply that

$$\sum_{s=1}^{\infty} |\phi_s^{(m)}| \leq \|\mathbf{M}'^{m-1}u\|_2 \|\mathbf{M}u\|_2 \sum_{s=1}^{\infty} \|\mathbf{M}(\mathbf{I}-uu')\|_2^{s-1} < \infty$$

for all $m \geq 1$. With the above in hand, we next show that $\lim_{h\to\infty} \beta_h^{\text{rx}} = 0$. Recall from Theorem 1 that the slope coefficient of the return predictability regression satisfies equation (13). Furthermore,

$$1 - \sum_{\tau=1}^{\infty} \delta^\tau \phi_\tau = 1 - \delta u'[\mathbf{I} - \delta\mathbf{M}(\mathbf{I}-uu')]^{-1}\mathbf{M}u = \frac{1}{1+\delta u'(\mathbf{I}-\delta\mathbf{M})^{-1}\mathbf{M}u} = \frac{1}{u'(\mathbf{I}-\delta\mathbf{M})^{-1}u},$$

where the second equality follows from the Sherman-Morrison formula and we are using ϕ_τ to denote $\phi_\tau^{(1)}$. Therefore, equation (13) implies that $\beta_h^{\text{rx}} = \delta u'(\mathbf{I}-\delta\mathbf{M})^{-1}u(\xi_h^* - \sum_{\tau=1}^\infty \phi_\tau \xi_{h-\tau}^*)$, and hence, by triangle inequality,

$$|\beta_h^{\text{rx}}| \le \delta |u'(\mathbf{I}-\delta\mathbf{M})^{-1}u| \left(|\xi_h^*| + \sum_{\tau=1}^{h} |\phi_\tau||\xi_{h-\tau}^*| + \sum_{\tau=1}^{\infty} |\phi_{\tau+h}||\xi_\tau^*| \right).$$

As a result,

$$\lim_{h\to\infty} |\beta_h^{\text{rx}}| \le \delta |u'(\mathbf{I}-\delta\mathbf{M})^{-1}u| \lim_{h\to\infty} \sum_{\tau=1}^{h} |\phi_\tau||\xi_{h-\tau}^*|, \tag{B.20}$$

where we are using $\lim_{h\to\infty}|\xi_h^*| = 0$ and the fact $\sum_{\tau=1}^\infty |\phi_\tau| < \infty$ implies $\lim_{h\to\infty}\sum_{\tau=1}^\infty |\phi_{\tau+h}||\xi_\tau^*| = 0$. To evaluate the right-hand side of equation (B.20), note that $\sum_{\tau=1}^h |\phi_\tau||\xi_{h-\tau}^*|$ is the hth term in the Cauchy product of the series $\sum_{\tau=1}^\infty |\phi_\tau|$ and $\sum_{\tau=0}^\infty |\xi_\tau^*|$. Since the first series is absolutely convergent, Merten's theorem guarantees that their Cauchy product also converges, thus implying that the right-hand side of equation (B.20) is equal to zero. A similar argument implies that $\lim_{h\to\infty}\beta_{h,m}^{\text{fe}} = 0$. □

Proof of Proposition 5

Lemma B.1. *If* $(\mathbf{M},u)$ *is the minimizer of equation (12), then* $\sum_{s=1}^\infty \phi_s \beta_{s,1}^{\text{fe}} = 0$, *where* $\phi_s = \phi_s^{(1)} = u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$.

Proof. The first-order condition of the optimization problem in equation (12) with respect to $\mathbf{M}'$ is given by $\sum_{s=1}^\infty(\xi_s^* - \sum_{\tau=1}^\infty \phi_\tau \xi_{s-\tau}^*)\mathrm{d}\phi_s/\mathrm{d}\mathbf{M}' = 0$, which in view of equation (14) can be rewritten as $\sum_{s=1}^\infty \beta_{s,1}^{\text{fe}}\mathrm{d}\phi_s/\mathrm{d}\mathbf{M}' = 0$. Differentiating $\phi_s = u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$ with respect to $\mathbf{M}'$ then implies that

$$\sum_{s=1}^{\infty}\left(uu'[\mathbf{M}(\mathbf{I}-uu')]^{s-1} + (\mathbf{I}-uu')\sum_{\tau=0}^{s-2}[\mathbf{M}(\mathbf{I}-uu')]^{\tau}\mathbf{M}uu'[\mathbf{M}(\mathbf{I}-uu')]^{s-\tau-2}\right)\beta_{s,1}^{\text{fe}} = 0.$$

Multiplying both sides of the above equation by u' from the left and $\mathbf{M}u$ from the right, using the fact that $u'u=1$, and noting that $\phi_s = u'[\mathbf{M}(\mathbf{I}-uu')]^{s-1}\mathbf{M}u$ then implies that $\sum_{s=1}^\infty \phi_s\beta_{s,1}^{\text{fe}} = 0$. □

With Lemma B.1 in hand, we now proceed to prove Proposition 5. Recall from the proof of Theorem 1 that the KL divergence of agents' model from the true data-generating process is given by equation (B.14). Furthermore, since the fundamental is normally distributed, we have

$$\mathbb{E}^*[\log f^*(x_{t+1}|x_t,\ldots)] = -\log\sqrt{2\pi} - \frac{1}{2}\mathbb{E}^*[\log \mathrm{var}_t^*(x_{t+1})] - \frac{1}{2}\mathbb{E}^*\left[\frac{(x_{t+1}-\mathbb{E}_t^*[x_{t+1}])^2}{\mathrm{var}_t^*(x_{t+1})}\right].$$

Note that $\text{var}_t^*(x_{t+1}) = \mathbb{E}_t^*[(x_{t+1} - \mathbb{E}_t^*[x_{t+1}])^2] = \mathbb{E}^*[(x_{t+1} - \mathbb{E}_t^*[x_{t+1}])^2]$. Therefore,

$$\begin{aligned}\mathbb{E}^*[\log f^*(x_{t+1}|x_t,\ldots)] &= -\log\sqrt{2\pi e} - \frac{1}{2}\log\mathbb{E}^*\left[\left(x_{t+1} - \sum_{\tau=1}^{\infty}\phi_\tau^* x_{t+1-\tau}\right)^2\right] \\ &= -\log\sqrt{2\pi e} - \frac{1}{2}\log\mathbb{E}^*[x_t^2] \\ &\quad -\frac{1}{2}\log\left(1 - 2\sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^* + \sum_{\tau=1}^{\infty}\sum_{s=1}^{\infty}\phi_s^*\phi_\tau^*\xi_{\tau-s}^*\right),\end{aligned}$$

where $(\phi_1^*, \phi_2^*, \ldots)$ are the coefficients of the autoregressive representation of the underlying data-generating process and ξ_τ^* is the autocorrelation of the fundamental at lag τ. Plugging the above into equation (B.14) therefore implies that

$$\begin{aligned}\text{KL}(\theta^*\|\theta) &= \frac{1}{2}\log\left(1 - 2\sum_{s=1}^{\infty}\phi_s\xi_s^* + \sum_{\tau=1}^{\infty}\phi_s\phi_\tau\xi_{\tau-s}^*\right) \\ &\quad -\frac{1}{2}\log\left(1 - 2\sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^* + \sum_{\tau=1}^{\infty}\sum_{s=1}^{\infty}\phi_s^*\phi_\tau^*\xi_{\tau-s}^*\right).\end{aligned}$$

We make three observations. First, equation (14) implies that $\sum_{s=1}^{\infty}\phi_s\xi_s^* - \sum_{s,\tau=1}^{\infty}\phi_s\phi_\tau\xi_{s-\tau}^* = \sum_{s=1}^{\infty}\phi_s\beta_{s,1}^{\text{fe}} = 0$, where the second equality follows from Lemma B.1. Therefore,

$$\text{KL}(\theta^*\|\theta) = \frac{1}{2}\log\left(1 - \sum_{\tau=1}^{\infty}\phi_\tau\xi_\tau^*\right) - \frac{1}{2}\log\left(1 - 2\sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^* + \sum_{\tau=1}^{\infty}\sum_{s=1}^{\infty}\phi_s^*\phi_\tau^*\xi_{\tau-s}^*\right).$$

Second, the definition of $\bar{\beta}^{\text{fe}}$ and equation (14) imply that $\bar{\beta}^{\text{fe}} = \sum_{h=1}^{\infty}\phi_h^*\xi_h^* - \sum_{h,\tau=1}^{\infty}\phi_h^*\phi_\tau\xi_{h-\tau}^*$. Hence,

$$\begin{aligned}\text{KL}(\theta^*\|\theta) &= \frac{1}{2}\log\left(\bar{\beta}^{\text{fe}} + 1 - \sum_{h=1}^{\infty}\phi_h^*\xi_h^* - \sum_{\tau=1}^{\infty}\phi_\tau\left(\xi_\tau^* - \sum_{h=1}^{\infty}\phi_h^*\xi_{h-\tau}^*\right)\right) \\ &\quad -\frac{1}{2}\log\left(1 - 2\sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^* + \sum_{\tau,s=1}^{\infty}\phi_s^*\phi_\tau^*\xi_{\tau-s}^*\right).\end{aligned}$$

Finally, note that $(\phi_1^*, \phi_2^*, \ldots)$ are the coefficients of the autoregressive representation of the true process: $\mathbb{E}_{t-1}^*[x_t] = \sum_{h=1}^{\infty}\phi_h^* x_{t-h}$. Therefore, $\xi_\tau^* = \sum_{h=1}^{\infty}\phi_h^*\xi_{h-\tau}^*$, and as a result,

$$\text{KL}(\theta^*\|\theta) = \frac{1}{2}\log\left(\bar{\beta}^{\text{fe}} + 1 - \sum_{h=1}^{\infty}\phi_h^*\xi_h^*\right) - \frac{1}{2}\log\left(1 - \sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^*\right).$$

Hence,

$$\bar{\beta}^{\text{fe}} = \left(1 - \sum_{\tau=1}^{\infty}\phi_\tau^*\xi_\tau^*\right)\left(e^{2\text{KL}(\theta^*\|\theta)} - 1\right). \tag{B.21}$$

This expression now establishes the claims in Proposition 5. First, since $\text{KL}(\theta^*\|\theta) \geq 0$, it follows that $\bar{\beta}^{\text{fe}} \geq 0$. Second, when $k \geq n$, we have $\text{KL}(\theta^*\|\theta) = 0$, which guarantees that $\bar{\beta}^{\text{fe}} = 0$.

Finally, since the first term on the right-hand side of equation (B.21) is independent of agents' model, it follows that minimizing $\mathrm{KL}(\theta^*\|\theta)$ is equivalent to minimizing $\bar{\beta}^{\mathrm{fe}}$. Therefore, Proposition 2 implies that increasing the number of factors in agents' model can only result in a reduction in $\bar{\beta}^{\mathrm{fe}}$. □

Proof of Proposition 6

By Theorem 1, the slope coefficients of the return and forecast-error predictability regressions are given by equations (13) and (14), where $\mathbf{M}$ and u are, respectively, a k-dimensional stable matrix and a k-dimensional unit vector, which minimize equation (12). When $k=1$, $\mathbf{M}$ is a scalar, denoted by μ, satisfying $|\mu|<1$ and $u\in\{-1,1\}$. This implies that $\phi_1^{(m)}=\mu^m$ and $\phi_s^{(m)}=0$ for all $m\geq 1$ and all $s\geq 2$. As a result, equations (13) and (14) reduce to

$$\beta_h^{\mathrm{rx}}=\frac{\delta}{1-\delta\mu}\left(\xi_h^*-\mu\xi_{h-1}^*\right)\quad\text{and}\quad\beta_{h,m}^{\mathrm{fe}}=\xi_{h+m-1}^*-\mu^m\xi_{h-1}^*,\tag{B.22}$$

respectively, and the objective function in equation (12) is given by

$$H(\mu,u)=1-2\mu\xi_1^*+\mu^2.\tag{B.23}$$

Optimizing the above over $\mu\in(-1,1)$ and plugging the result into equation (B.22) establishes the result. □

Proof of Proposition 7

Let $p(t)=t^n-\sum_{i=1}^n\gamma_i t^{n-i}$ denote the characteristic polynomial of $\mathbf{A}^*$, which is the n-dimensional square matrix that governs the transitional dynamics of the true data-generating process. By the Cayley-Hamilton theorem (Horn and Johnson, 2013, p. 109), matrix $\mathbf{A}^*$ satisfies its characteristic polynomial, that is, $\mathbf{A}^{*n}=\sum_{i=1}^n\gamma_i\mathbf{A}^{*n-i}$. Multiplying both sides of this equation by $\mathbf{A}^{*h-n}$ implies that

$$\mathbf{A}^{*h}=\gamma_1\mathbf{A}^{*h-1}+\cdots+\gamma_n\mathbf{A}^{*h-n}\quad\text{for all } h\geq n.\tag{B.24}$$

It is easy to verify that the auto-covariance of the true process at lag h is given by $\mathbb{E}^*[x_tx_{t+h}]=c^{*\prime}\mathbf{A}^{*h}\mathbf{Z}c^*$, where $\mathbf{Z}=\sum_{j=0}^{\infty}(\mathbf{A}^{*j}\mathbf{B}^*)(\mathbf{A}^{*j}\mathbf{B}^*)'$. Therefore, multiplying both sides of equation (B.24) by $c^{*\prime}$ from the left and by $\mathbf{Z}c^*$ from the right and dividing by $\mathbb{E}^*[x_t^2]$ implies that

$$\xi_h^*=\gamma_1\xi_{h-1}^*+\cdots+\gamma_n\xi_{h-n}^*\quad\text{for all } h\geq n.\tag{B.25}$$

Recall from equation (15) that when $k=1$, the slope coefficient of the return predictability regression is given by $\beta_h^{\mathrm{rx}}=\delta(\xi_h^*-\xi_1^*\xi_{h-1}^*)/(1-\delta\xi_1^*)$. This observation, combined with equation (B.25), thus implies that

$$\beta_h^{\mathrm{rx}}=\gamma_1\beta_{h-1}^{\mathrm{rx}}+\cdots+\gamma_n\beta_{h-n}^{\mathrm{rx}}\quad\text{for all } h\geq n+1.\tag{B.26}$$

Equation (B.26) establishes that β_h^{rx} satisfies a recursive equation with coefficients given by the coefficients of the characteristic polynomial of matrix $\mathbf{A}^*$. Since the roots of the characteristic polynomial of $\mathbf{A}^*$ are equal to its eigenvalues, the solution to the recursive equation in equation

(B.26) when all eigenvalues of $\mathbf{A}^*$ are distinct is given by

$$\beta_h^{\text{rx}} = \sum_{i=1}^{n} \zeta_i \lambda_i^{h-1}$$

for all h, where $(\lambda_1, \dots, \lambda_n)$ denote the eigenvalues of $\mathbf{A}^*$ and $(\zeta_1, \dots, \zeta_n)$ are a collection of constants that do not depend on h. Furthermore, note that, by equation (15), $\beta_1^{\text{rx}} = 0$, which implies that $\sum_{i=1}^{n} \zeta_i = 0$. Therefore, when $n = 2$, the slope coefficients of the predictability regression (6) are given by $\beta_h^{\text{rx}} = \zeta_1(\lambda_1^{h-1} - \lambda_2^{h-1})$ for all $h \geq 1$, where ζ_1 is some constant and λ_1 and λ_2 are the eigenvalues of $\mathbf{A}^*$. Consequently,

$$\beta_h^{\text{rx}} = \beta_2^{\text{rx}} \left(\frac{\lambda_1^{h-1} - \lambda_2^{h-1}}{\lambda_1 - \lambda_2} \right). \tag{B.27}$$

With the above in hand, we can now prove the various statements of the proposition. In what follows, we assume that λ_1 is the eigenvalue with the largest modulus, that is, $|\lambda_1| \geq |\lambda_2|$.

Proof of Part 1. Suppose both λ_1 and λ_2 are real and that $\lambda_1 > 0$. It is immediate that the right-hand side of equation (B.27) never changes sign. Therefore, β_h^{rx} has the same sign as β_2^{rx} for all h. □

Proof of Part 2. Suppose both λ_1 and λ_2 are real and that $\lambda_1 < 0$. In this case, the right-hand side of equation (B.27) has the same sign as β_2^{rx} when h is even, whereas it has the opposite sign of β_2^{rx} when h is odd. This means that β_h^{rx} changes sign at every horizon. □

Proof of Part 3. Suppose both λ_1 and λ_2 are complex. Since they are eigenvalues of real matrix $\mathbf{A}^*$, they form a complex conjugate pair. We can therefore represent them in polar coordinates as $\lambda_1 = |\lambda| e^{i\varphi}$ and $\lambda_2 = |\lambda| e^{-i\varphi}$, where without loss of generality we assume that $0 < \varphi < \pi$. Plugging these values into the right-hand side of equation (B.27) implies that

$$\beta_{h+1}^{\text{rx}} = \beta_2^{\text{rx}} |\lambda|^{h-1} \left(\frac{e^{ih\varphi} - e^{-ih\varphi}}{e^{i\varphi} - e^{-i\varphi}} \right) = \beta_2^{\text{rx}} |\lambda|^{h-1} \frac{\sin(h\varphi)}{\sin\varphi}$$

for all $h \geq 1$. Since $0 < \varphi < \pi$, the denominator of the ratio on the right-hand side of the above equation is always positive. Therefore, β_{h+1}^{rx} has the same sign as β_2^{rx} if $\sin(h\varphi) > 0$ and has the opposite sign if β_2^{rx} if $\sin(h\varphi) < 0$. Therefore, β_{h+1}^{rx} changes sign at frequency π/φ. □

Proof of Proposition 8

Recall from the proof of Theorem 1 that any k-factor model $\theta \in \Theta_k$ can be represented by the pair $(\mathbf{M}, u)$ defined in equation (B.11). Also recall from equation (B.12) that agents' m-step-ahead forecasts are given by

$$\mathbb{E}_t[x_{t+m}] = \sum_{\tau=1}^{\infty} u' \mathbf{M}^{m-1} [\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M} u x_{t+1-\tau}.$$

Comparing this equation to the representation in equation (17) implies that coefficients $(\phi_1^{(m)}, \phi_2^{(m)}, \ldots)$ can be expressed in terms of $\mathbf{M}$ and u as

$$\phi_\tau^{(m)} = u'\mathbf{M}^{m-1}\mathbf{S}^{\tau-1}\mathbf{M}u \quad \text{for all } \tau, m \geq 1, \tag{B.28}$$

where $\mathbf{S} = \mathbf{M}(\mathbf{I} - uu')$. As a result, the elements of Hankel matrix $\mathbf{\Phi}^{(m)}$ in equation (18) are given by $\mathbf{\Phi}_{ij}^{(m)} = u'\mathbf{M}^{m-1}\mathbf{S}^{i+j-2}\mathbf{M}u$ for all $i, j \geq 1$. In what follows, we first show that $\text{rank}(\mathbf{\Phi}^{(m)}) = k$ for $m = 1$ and then extend the result to all $m \geq 2$.

Set $m = 1$. In view of equation (B.28), we can express $\mathbf{\Phi}^{(1)}$ as the product two matrices as follows:

$$\mathbf{\Phi}^{(1)} = \mathbf{OC}, \tag{B.29}$$

where $\mathbf{O} \in \mathbb{R}^{\infty \times k}$ and $\mathbf{C} \in \mathbb{R}^{k \times \infty}$ are given by

$$\mathbf{O} = \begin{bmatrix} u' \\ u'\mathbf{S} \\ u'\mathbf{S}^2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \mathbf{C} = \begin{bmatrix} \mathbf{M}u & \mathbf{SM}u & \mathbf{S}^2\mathbf{M}u & \cdots \end{bmatrix},$$

respectively. The decomposition in equation (B.29), together with Sylvester's rank inequality (Horn and Johnson, 2013, p. 13), implies that

$$\text{rank}(\mathbf{O}) + \text{rank}(\mathbf{C}) - k \leq \text{rank}(\mathbf{\Phi}^{(1)}) \leq \min\{\text{rank}(\mathbf{O}), \text{rank}(\mathbf{C})\}.$$

Given these inequalities, it is immediate that if $\text{rank}(\mathbf{O}) = \text{rank}(\mathbf{C}) = k$, then $\text{rank}(\mathbf{\Phi}^{(1)}) = k$, which establishes the result. Therefore, it is sufficient to show that $\text{rank}(\mathbf{O}) = \text{rank}(\mathbf{C}) = k$.

We first show that, generically, $\text{rank}(\mathbf{O}) = k$. Suppose to the contrary that $\text{rank}(\mathbf{O}) < k$, which implies that the pair $(u', \mathbf{S})$ is unobservable. Hence, by Popov–Belevitch–Hautus (PBH) observability criterion (Hespanha, 2018, p. 191), there exist a non-zero vector w and a scalar λ such that $\mathbf{S}w = \lambda w$ and $u'w = 0$. Since $\mathbf{S} = \mathbf{M}(\mathbf{I} - uu')$, it must be the case that $\mathbf{M}w = \lambda w$. This means that the pair $(\mathbf{M}, u)$ is such that a right eigenvector of $\mathbf{M}$ is orthogonal to u, a statement that does not hold generically. Therefore, it must be the case that $\text{rank}(\mathbf{O}) = k$.

Next, we establish that, generically, $\text{rank}(\mathbf{C}) = k$. Suppose to the contrary that $\text{rank}(\mathbf{C}) < k$, which implies that the pair $(\mathbf{S}, \mathbf{M}u)$ is uncontrollable. Hence, by the PBH criterion for controllability (Hespanha, 2018, p. 152), there exist a non-zero vector w and a scalar λ such that $w'\mathbf{S} = \lambda w'$ and $w'\mathbf{M}u = 0$. Since $\mathbf{S} = \mathbf{M}(\mathbf{I} - uu')$, it must be the case that

$$w'\mathbf{M} = \lambda w', \quad \lambda w'u = 0.$$

The above conditions can hold only if either (i) matrix $\mathbf{M}$ is rank deficient (with an eigenvalue $\lambda = 0$) or (ii) the pair $(\mathbf{M}, u)$ is such that a left eigenvector of $\mathbf{M}$ is orthogonal to u. But neither of these conditions can hold generically. Therefore, for a generic $\theta \in \Theta_k$, it must be the case that $\text{rank}(\mathbf{C}) = k$. This, together with the fact that $\text{rank}(\mathbf{O}) = k$, establishes that $\text{rank}(\mathbf{\Phi}^{(1)}) = k$ generically.

We now show that $\text{rank}(\mathbf{\Phi}^{(m)}) = k$ for $m \geq 2$. Start with the observation that the expression in equation (B.28) implies that $\phi_\tau^{(m)} = \phi_{\tau+1}^{(m-1)} + \phi_1^{(m-1)}\phi_\tau^{(1)}$. Iterating on this equation, we get

$$\phi_\tau^{(m)} = \sum_{r=0}^{m-1} \phi_1^{(m-r-1)}\phi_{\tau+r}^{(1)},$$

with the convention that $\phi_1^{(0)} = 1$. Since $\boldsymbol{\Phi}_{i,j}^{(m)} = \phi_{i+j-1}^{(m)}$, we have

$$\boldsymbol{\Phi}_{i,j}^{(m)} = \sum_{r=0}^{m-1} \phi_1^{(m-r-1)} \boldsymbol{\Phi}_{i+r,j}^{(1)} \quad \text{for all } m \geq 2 \text{ and all } i,\, j \geq 1.$$

Therefore, $\boldsymbol{\Phi}^{(m)} = \mathbf{L}^{(m)} \boldsymbol{\Phi}^{(1)}$ for an upper-triangular matrix $\mathbf{L}^{(m)}$ whose diagonal elements are all equal to $\phi_1^{(m-1)}$. Hence, $\mathrm{rank}(\boldsymbol{\Phi}^{(m)}) = \mathrm{rank}(\boldsymbol{\Phi}^{(1)}) = k$ as long as $\phi_1^{(m-1)} \neq 0$. But note that $\phi_1^{(m-1)} = u'\mathbf{M}^{m-1}u$, which is different from zero generically. Thus, generically, $\mathrm{rank}(\boldsymbol{\Phi}^{(m)}) = k$ for all $m \geq 2$. □

Proof of Proposition A.1

We first establish that an agent who uses a k-factor model to represent the underlying data-generating process makes forecasts about the future by relying on k moving averages of her past observations, one of which is the most recent realization of the fundamental.

Consider an agent who uses a k-factor model with parameters $\theta = (\mathbf{A}, \mathbf{B}, c)$ to make predictions about future realizations of the fundamental. Let $\hat{\Sigma}$ and g be the corresponding solution to the algebraic Riccati equation and the Kalman gain given by equations (B.7) and (B.8), respectively. It is immediate to verify that $(\mathbf{A} - gc')\hat{\Sigma}c = 0$, which means that $\mathbf{A} - gc'$ is a rank-deficient matrix. Therefore, there exists an invertible matrix $\mathbf{V}$ such that the first row of the product $\mathbf{V}(\mathbf{A} - gc')$ is equal to zero, that is, $e_1'\mathbf{V}(\mathbf{A} - gc') = 0$, where e_1 denotes the first unit vector. Depending on properties of $\mathbf{V}$, we consider two separate cases.

First, suppose $\mathbf{V}$ is such that $e_1'\mathbf{V}g \neq 0$. Let $\mathbf{P} = \frac{1}{e_1'\mathbf{V}g}\mathbf{V}$ and consider the following vector of k moving averages of the realizations of the fundamental:

$$s_t = \mathbf{P}gx_t + \mathbf{P}(\mathbf{A} - gc')\mathbf{P}^{-1}s_{t-1}, \tag{B.30}$$

which can be rewritten as $s_t = \sum_{r=0}^{\infty}[\mathbf{P}(\mathbf{A} - gc')\mathbf{P}^{-1}]^r\mathbf{P}gx_{t-r} = \mathbf{P}\sum_{r=0}^{\infty}(\mathbf{A} - gc')^r gx_{t-r}$. But recall that the forecasts of an agent who uses a k-factor model with parameters $\theta = (\mathbf{A}, \mathbf{B}, c)$ are given by equation (B.9). As a result, one can write the agent's forecasts in terms of vector s_t:

$$\mathbb{E}_t[x_{t+\tau}] = c'\mathbf{P}^{-1}(\mathbf{P}\mathbf{A}\mathbf{P}^{-1})^{\tau-1}s_t. \tag{B.31}$$

Comparing equations (B.30)–(B.31) to equations (A.1)–(A.2) implies that an agent who uses a k-factor model can obtain the same exact forecasts by maintaining a vector of k moving averages of the form $s_{t+1} = wx_{t+1} + \mathbf{Q}s_t$, where $w = \mathbf{P}g$ and $\mathbf{Q} = \mathbf{P}(\mathbf{A} - gc')\mathbf{P}^{-1}$. The proof for this case is therefore complete once we show that one of the summary statistics that is maintained by the agent is equal to the most recent realization of the fundamental. To this end, note that

$$e_1's_t = e_1'wx_t + e_1'\mathbf{Q}s_{t-1} = e_1'\mathbf{P}gx_t + e_1'\mathbf{P}(\mathbf{A} - gc')\mathbf{P}^{-1}s_{t-1}.$$

By construction, $e_1'\mathbf{P}g = 1$ and $e_1'\mathbf{V}(\mathbf{A} - gc') = 0$. This therefore guarantees that $s_{1t} = x_t$.

As the second case, suppose matrix $\mathbf{V}$ is such that $e_1'\mathbf{V}g = 0$ and consider the following vector of k moving averages of the realizations of the fundamental:

$$s_t = (\mathbf{V}g + e_1)x_t + \mathbf{V}(\mathbf{A} - gc')\mathbf{V}^{-1}(\mathbf{I} - e_1e_1')s_{t-1}. \tag{B.32}$$

The assumption that $e_1'\mathbf{V}g = 0$ together with $e_1'\mathbf{V}(\mathbf{A} - gc') = 0$ implies that $e_1's_t = x_t$. Therefore, the first element of s_t is equal to the most recent realization of the fundamental.

Additionally, note that the vector of moving averages in equation (B.32) can be rewritten as

$$
\begin{aligned}
s_t &= \sum_{r=0}^{\infty}[\mathbf{V}(\mathbf{A}-gc')\mathbf{V}^{-1}(\mathbf{I}-e_1e_1')]^r(\mathbf{V}g+e_1)x_{t-r}\\
&= \sum_{r=0}^{\infty}\mathbf{V}(\mathbf{A}-gc')^r\mathbf{V}^{-1}(\mathbf{I}-e_1e_1')(\mathbf{V}g+e_1)x_{t-r},
\end{aligned}
$$

where the second equality follows from the fact that $e_1'\mathbf{V}(\mathbf{A}-gc')=0$. Using $e_1'\mathbf{V}g=0$ one more time implies that $s_t=\mathbf{V}\sum_{r=0}^{\infty}(\mathbf{A}-gc')^r gx_{t-r}$. Multiplying both sides of this equation by $(\mathbf{I}-e_1e_1')$ from the left implies that

$$
(\mathbf{I}-e_1e_1')s_t=\mathbf{V}\sum_{r=0}^{\infty}(\mathbf{A}-gc')^r gx_{t-r}-e_1e_1'\mathbf{V}gx_t=\mathbf{V}\sum_{r=0}^{\infty}(\mathbf{A}-gc')^r gx_{t-r}, \tag{B.33}
$$

where the first equality follows from the fact that $e_1'\mathbf{V}(\mathbf{A}-gc')^r=0$ and the second equality is a consequence of $e_1'\mathbf{V}g=0$. Additionally, recall that the forecasts of an agent who uses a k-factor model with parameters $\theta=(\mathbf{A},\mathbf{B},c)$ are given by (B.9). Therefore, replacing for $\sum_{r=0}^{\infty}(\mathbf{A}-gc')^r gx_{t-r}$ with $\mathbf{V}^{-1}(\mathbf{I}-e_1e_1')s_t$ from equation (B.33) into equation (B.9) implies that the agent's forecasts can be written in terms of the vector of moving averages s_t in equation (B.32):

$$
\begin{aligned}
\mathbb{E}_t[x_{t+\tau}] &= c'\mathbf{V}^{-1}(\mathbf{VAV}^{-1})^{\tau-1}(\mathbf{I}-e_1e_1')s_t\\
&= c'\mathbf{V}^{-1}\left[(\mathbf{I}-e_1e_1')\left(e_1c'\mathbf{V}^{-1}+\mathbf{VAV}^{-1}\right)\right]^{\tau-1}(\mathbf{I}-e_1e_1')s_t,
\end{aligned}
$$

where the second equality follows form the fact that $e_1'\mathbf{VA}=e_1'\mathbf{V}gc'=0$. Hence,

$$
\mathbb{E}_t[x_{t+\tau}]=c'\mathbf{V}^{-1}(\mathbf{I}-e_1e_1')\left[\left(e_1c'\mathbf{V}^{-1}+\mathbf{VAV}^{-1}\right)(\mathbf{I}-e_1e_1')\right]^{\tau-1}s_t. \tag{B.34}
$$

Comparing equations (B.32) and (B.34) with equations (A.1) and (A.2) once again establishes that an agent who uses a k-factor model can obtain the same exact forecasts by maintaining a vector of k moving averages of the form $s_{t+1}=wx_{t+1}+\mathbf{Q}s_t$, where $w=\mathbf{V}g+e_1$ and $\mathbf{Q}=\mathbf{V}(\mathbf{A}-gc')\mathbf{V}^{-1}(\mathbf{I}-e_1e_1')$. Together, the two cases therefore imply that irrespective of the value of $e_1'\mathbf{V}g$, the agent using a k-factor model can always be represented as an agent with limited memory as formalized by equations (A.1) and (A.2).

Next, we turn to proving the converse implication. Consider an agent who maintains k moving averages of the past realizations of the fundamental according to equation (A.1) and makes forecasts according to equation (A.2). Rewriting equations (A.1) and (A.2) in matrix form, we get $s_{t+1}=wx_{t+1}+\mathbf{Q}s_t$ and $\mathbb{E}_t[x_{t+\tau}]=v_\tau's_t$, where s_t is the vector of k moving averages at time t, w and v_τ are k-dimensional vectors, and $\mathbf{Q}$ is a square $k\times k$ matrix. Multiplying both sides of $s_{t+1}=wx_{t+1}+\mathbf{Q}s_t$ by v_τ' from the left and using the assumption that $\mathbb{E}_{t+1}[x_{t+1+\tau}]=v_\tau's_{t+1}$ implies that

$$
\mathbb{E}_{t+1}[x_{t+1+\tau}]=v_\tau's_{t+1}=v_\tau'wx_{t+1}+v_\tau'\mathbf{Q}s_t.
$$

Taking conditional expectations from both sides and using the assumption that the agent's expectations satisfy the law of iterated expectations implies that $\mathbb{E}_t[x_{t+1+\tau}]=v_\tau'w\mathbb{E}_t[x_{t+1}]+v_\tau'\mathbf{Q}s_t$. Note that $\mathbb{E}_t[x_{t+1+\tau}]=v_{\tau+1}'s_t$ and $\mathbb{E}_t[x_{t+1}]=v_1's_t$. Hence, $v_{\tau+1}'=v_\tau'(wv_1'+\mathbf{Q})$ for all $\tau\geq 1$, which in turn implies that $v_\tau'=v_1'(wv_1'+\mathbf{Q})^{\tau-1}$. Therefore, the forecast of the agent who uses

k moving averages for the realization of the fundamental can be expressed as

$$\mathbb{E}_t[x_{t+\tau}] = v_1'(wv_1' + \mathbf{Q})^{\tau-1} s_t = v_1'(wv_1' + \mathbf{Q})^{\tau-1} \sum_{r=0}^{\infty} \mathbf{Q}^r w x_{t-r} \tag{B.35}$$

for all $\tau \geq 1$. Additionally, note that the assumption that $s_{1t} = x_t$ implies that $e_1'w = 1$ and $e_1'\mathbf{Q} = 0$, which means that $\mathbf{Q}$ is a rank-deficient matrix.

We now show that the forecasts in equation (B.35) coincide with forecasts of an agent who assumes the true data-generating process is a k-factor model with parameters $\theta = (\mathbf{A}, \mathbf{B}, c)$, where

$$\mathbf{A} = wv_1' + \mathbf{Q}, \quad \mathbf{B} = pe_1', \quad c = v_1, \tag{B.36}$$

in which e_1 is the first unit vector and p is a vector in the right null space of $\mathbf{Q}$, that is, $\mathbf{Q}p = 0$. Given model parameters in equation (B.36), it follows from equation (B.8) that the corresponding Kalman gain is given by

$$g = (wv_1' + \mathbf{Q})\hat{\Sigma} v_1 (v_1' \hat{\Sigma} v_1)^{-1}, \tag{B.37}$$

where $\hat{\Sigma}$ is the solution to the algebraic Riccati equation in equation (B.7) and satisfies

$$\begin{aligned}\hat{\Sigma} &= (wv_1' + \mathbf{Q})\left(\hat{\Sigma} - \frac{1}{v_1'\hat{\Sigma}v_1}\hat{\Sigma}v_1v_1'\hat{\Sigma}\right)(v_1 w' + \mathbf{Q}') + pp' \\ &= \mathbf{Q}\left(\hat{\Sigma} - \frac{1}{v_1'\hat{\Sigma}v_1}\hat{\Sigma}v_1v_1'\hat{\Sigma}\right)\mathbf{Q}' + pp'.\end{aligned}$$

Recall that, by assumption, $\mathbf{Q}p = 0$. It is therefore immediate to verify that the positive semi-definite matrix $\hat{\Sigma} = pp'$ satisfies the above equation. Therefore, the Kalman gain in equation (B.37) is given by $g = w$. Replacing for g and the parameters in equation (B.36) into the expression for the agent's forecasts in equation (B.9) leads to

$$\mathbb{E}_t[x_{t+\tau}] = v_1'(wv_1' + \mathbf{Q})^{\tau-1} \sum_{r=0}^{\infty} \mathbf{Q}^r w x_{t-r},$$

which coincides with the expression in equation (B.35). In other words, the forecasts of the agent with the k-factor model with parameters in equation (B.36) coincide with those of the agent with k moving averages with parameters given by $(w, \mathbf{Q}, \{v_\tau\})$. □

Proof of Proposition A.2

Proof of Part 1. Fix the compact set $\widehat{\Theta}_k \subseteq \Theta_k$ and let $\hat{\theta}_t^{\mathrm{ML}} \in \arg\max_{\theta \in \widehat{\Theta}_k} f^\theta(x_t, x_{t-1}, \ldots, x_0)$ denote the agent's maximum likelihood estimator. Also define $h(\theta^*, \theta) = \mathbb{E}^*[\log f^\theta(x_{t+1}|x_t, x_{t-1}, \ldots)]$. Observe that $\mathrm{KL}(\theta^* \| \theta) = h(\theta^*, \theta^*) - h(\theta^*, \theta)$. As a result, $\arg\min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta) = \arg\max_{\theta \in \widehat{\Theta}_k} h(\theta^*, \theta)$, and in particular,

$$h(\theta^*, \theta^{\mathrm{KL}}) - h(\theta^*, \hat{\theta}_t^{\mathrm{ML}}) \geq 0 \tag{B.38}$$

for all $\theta^{\mathrm{KL}} \in \arg\min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta)$.

Next, note that

$$\begin{aligned}h(\theta^*,\theta^{\mathrm{KL}})-h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})=h(\theta^*,\theta^{\mathrm{KL}})&-\frac{1}{t}\log f^{\theta^{\mathrm{KL}}}(x_t,\ldots,x_0)\\&+\frac{1}{t}\log f^{\theta^{\mathrm{KL}}}(x_t,\ldots,x_0)-\frac{1}{t}\log f^{\hat{\theta}_t^{\mathrm{ML}}}(x_t,\ldots,x_0)\\&+\frac{1}{t}\log f^{\hat{\theta}_t^{\mathrm{ML}}}(x_t,\ldots,x_0)-h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})\end{aligned}$$

Since $\hat{\theta}_t^{\mathrm{ML}}=\arg\max_{\theta\in\widehat{\Theta}_k}f^{\theta}(x_t,x_{t-1},\ldots,x_0)$, it follows that

$$\begin{aligned}h(\theta^*,\theta^{\mathrm{KL}})-h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})&\leq h(\theta^*,\theta^{\mathrm{KL}})-\frac{1}{t}\log f^{\theta^{\mathrm{KL}}}(x_t,\ldots,x_0)\\&\quad+\frac{1}{t}\log f^{\hat{\theta}_t^{\mathrm{ML}}}(x_t,\ldots,x_0)-h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})\\&\leq 2\sup_{\theta\in\hat{\Theta}_k}\left|\frac{1}{t}\log f^{\theta}(x_t,\ldots,x_0)-h(\theta^*,\theta)\right|\end{aligned}$$

for all $\theta^{\mathrm{KL}}\in\arg\min_{\theta\in\widehat{\Theta}_k}\mathrm{KL}(\theta^*\|\theta)$. Taking limits as $t\to\infty$ from both sides of the above inequality and using the uniform law of large numbers (White, 1994, Theorem A.2.2) implies that

$$h(\theta^*,\theta^{\mathrm{KL}})-\lim_{t\to\infty}h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})\leq 0,\quad \mathbb{P}^*\text{-almost surely.}$$

This inequality, together with the inequality in equation (B.38) therefore guarantees that

$$\lim_{t\to\infty}h(\theta^*,\hat{\theta}_t^{\mathrm{ML}})=h(\theta^*,\theta^{\mathrm{KL}}),\quad \mathbb{P}^*\text{-almost surely.}$$

Subtracting $h(\theta^*,\theta^*)$ from both sides of the above equation and using the fact that $\mathrm{KL}(\theta^*\|\theta)=h(\theta^*,\theta^*)-h(\theta^*,\theta)$ then establishes the result. □

Proof of Part 2. Fix an arbitrary compact subset $\widehat{\Theta}_k\subseteq\Theta_k$ of the class of k-factor models and let $\hat{\theta}_t^{\mathrm{MSE}}\in\arg\min_{\theta\in\widehat{\Theta}_k}\frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1}-\mathbb{E}_\tau^\theta[x_{\tau+1}])^2$ be a model that minimizes the time-series average of agents' forecast errors squared given observations from time 0 to time t. Also let $\mathrm{MSE}(\theta^*,\theta)=\mathbb{E}^*[x_{t+1}-\mathbb{E}_t^\theta[x_{t+1}]]^2$ denote the *expected* forecast error squared when agents forecast the future realizations of the fundamental using model θ but the true model is θ^*. Finally, let $\theta^{\mathrm{KL}}\in\arg\min_{\theta\in\widehat{\Theta}_k}\mathrm{KL}(\theta^*\|\theta)$ denote a model with minimum KL divergence from the true process, θ^*.

As a first step, we establish that

$$\mathrm{MSE}(\theta^*,\theta)-\mathrm{MSE}(\theta^*,\theta^{\mathrm{KL}})\geq 0 \tag{B.39}$$

for all $\theta\in\widehat{\Theta}_k$. To this end, recall from the proof of Theorem 1 that any model $\theta\in\Theta_k$ can be equivalently represented by the tuple $(\mathbf{M},u,\hat{\sigma}_x^{-2})$, where $\mathbf{M}$ is a square k-dimensional stable matrix, u is a k-dimensional vector of unit length, and $\hat{\sigma}_x^2$ is a positive scalar. We also established that, given agents' model, their forecasts of the future realizations of the fundamental are given by equation (B.12) for all $s \geq 1$. Therefore, $\mathbb{E}_t[x_{t+1}] = \sum_{\tau=1}^{\infty} \phi_\tau x_{t+1-\tau}$, where $\phi_\tau = u'[\mathbf{M}(\mathbf{I} - uu')]^{\tau-1}\mathbf{M}u$. As a result,

$$\mathrm{MSE}(\theta^*, \theta) = \mathbb{E}^*\left[\left(x_{t+1} - \sum_{\tau=1}^{\infty} \phi_\tau x_{t+1-\tau}\right)^2\right] = \mathbb{E}^*[x_{t+1}^2] H(\mathbf{M}, u), \tag{B.40}$$

where $H(\mathbf{M}, u)$ is given by equation (12). Since $\mathbb{E}^*[x_{t+1}^2]$ only depends on the true data-generating process and is independent of agents' model θ, the above equation implies that minimizing $\mathrm{MSE}(\theta^*, \theta)$ is equivalent to minimizing $H(\mathbf{M}, u)$ over $\mathbf{M}$ and u. But recall from the proof of Theorem 1 that minimizing $\mathrm{KL}(\theta^* || \theta)$ over the set of k-factor models is also equivalent to minimizing equation (12) with respect to $\mathbf{M}$ and u, that is, $\arg\min_{\theta \in \widehat{\Theta}_k} \mathrm{MSE}(\theta^*, \theta) = \arg\min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* || \theta)$. This establishes equation (B.39).

Next, note that

$$\begin{aligned}\mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}) = {} & \mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\hat{\theta}_t^{\mathrm{MSE}}}[x_{\tau+1}])^2 \\ & + \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\hat{\theta}_t^{\mathrm{MSE}}}[x_{\tau+1}])^2 - \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\theta^{\mathrm{KL}}}[x_{\tau+1}])^2 \\ & + \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\theta^{\mathrm{KL}}}[x_{\tau+1}])^2 - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}).\end{aligned}$$

Since $\hat{\theta}_t^{\mathrm{MSE}} \in \arg\min_{\theta \in \widehat{\Theta}_k} \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\theta}[x_{\tau+1}])^2$, it follows that

$$\begin{aligned}\mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}) \leq {} & \mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\hat{\theta}_t^{\mathrm{MSE}}}[x_{\tau+1}])^2 \\ & + \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\theta^{\mathrm{KL}}}[x_{\tau+1}])^2 - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}),\end{aligned}$$

and as a result,

$$\mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}) \leq 2\sup_{\theta \in \widehat{\Theta}_k}\left|\mathrm{MSE}(\theta^*, \theta) - \frac{1}{t}\sum_{\tau=0}^{t}(x_{\tau+1} - \mathbb{E}_\tau^{\theta}[x_{\tau+1}])^2\right|.$$

Taking limits as $t \to \infty$ from both sides of the above inequality and using the uniform law of large numbers (White, 1994, Theorem A.2.2) implies that $\lim_{t\to\infty} \mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}) \leq 0$ with $\mathbb{P}^*$-probability one. Furthermore, note that equation (B.39) guarantees that $\lim_{t\to\infty} \mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) - \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}) \geq 0$. Therefore,

$$\lim_{t\to\infty} \mathrm{MSE}(\theta^*, \hat{\theta}_t^{\mathrm{MSE}}) = \mathrm{MSE}(\theta^*, \theta^{\mathrm{KL}}), \quad \mathbb{P}^*\text{-almost surely.}$$

Dividing both sides by $\mathbb{E}^*[x_{t+1}^2]$ and using equations (B.40) and (B.14) establishes the result. □

Proof of Proposition A.3

Since rational agents can fully construct the model used by behavioural agents, the regress of expectations in equation (A.4) is given by

$$\bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} \cdots \bar{\mathbb{E}}_{t+\tau}[x_{t+\tau}] = \gamma^{\tau+1} \mathbb{E}_t^*[x_{t+\tau}] + (1-\gamma) \sum_{s=0}^{\tau} \gamma^s \mathbb{E}_t^* \mathbb{E}_{t+s}[x_{t+\tau}]$$

for all $\tau \geq 0$, where $\mathbb{E}[\cdot]$ and $\mathbb{E}^*[\cdot]$ are the subjective expectations of the behavioural and rational agents, respectively, and γ denotes the fraction of rational agents. Therefore, the excess return in the economy with a fraction γ of rational agents is given by

$$\begin{aligned} \mathrm{rx}_t(\gamma) = (1-\gamma) \sum_{s=0}^{\infty} (\delta\gamma)^s \sum_{\tau=1}^{\infty} \delta^\tau \Big(\mathbb{E}_t^* \mathbb{E}_{t+s}[x_{t+\tau+s-1}] - \mathbb{E}_{t-1}^* \mathbb{E}_{t+s-1}[x_{t+\tau+s-1}] \Big) \\ + \sum_{\tau=1}^{\infty} (\delta\gamma)^\tau \Big(\mathbb{E}_t^*[x_{t+\tau-1}] - \mathbb{E}_{t-1}^*[x_{t+\tau-1}] \Big). \end{aligned}$$

Taking expectations from both sides of the above equation and using the expression in equation (5) for excess returns in the representative-agent economy therefore implies that

$$\mathbb{E}_{t-1}^*[\mathrm{rx}_t(\gamma)] = (1-\gamma) \sum_{s=0}^{\infty} (\delta\gamma)^s \mathbb{E}_{t-1}^*[\mathrm{rx}_{t+s}(0)].$$

Therefore, expected excess return in the heterogenous-agent economy is the discounted sum of all future expected excess returns of a representative-agent economy populated only by the behavioural agents. Multiplying both sides of the above equation by x_{t-h} and taking expectations $\mathbb{E}^*[\cdot]$ then establishes the result. □

Acknowledgments. We are grateful to the editor and three anonymous referees for helpful comments and suggestions. We also thank Fernando Alvarez, Turan Bali, Pedro Bordalo, Andrea Buffa, Julien Cujean, Charles Engel, Ignacio Esponda, Xavier Gabaix, Tarek Hassan, Ben Hébert, Ali Jadbabaie, Lars Hansen, Jen-Jen La'O, Xinyang Li, Stefan Nagel, Stavros Panageas, Carolin Pflueger, Farzad Saidi, Harald Uhlig, Dimitri Vayanos, Mike Woodford, Hao Xing, and numerous seminar and conference participants.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

Data Availability Statement

The replication package is available on Zenodo at https://doi.org/10.5281/zenodo.7843956. The raw data on interest rate forecasts were provided by Consensus Economics Inc. under license. Details on the data sources, code, and computational requirements can be found in the ReadMe file in the replication package.

REFERENCES

ALLEN, F., MORRIS, S. and SHIN, H. S. (2006), "Beauty Contests and Iterated Expectations in Asset Markets", *Review of Financial Studies*, **19**, 719–752.

AL'PIN, Y. A. (2017), "The Hankel Matrix Rank Theorem Revisited", *Linear Algebra and its Applications*, **534**, 97–101.

ANGELETOS, G.-M. and HUO, Z. (2021), "Myopia and Anchoring", *American Economic Review*, **111**, 1166–1200.

ANGELETOS, G.-M., HUO, Z. and SASTRY, K. A. (2020), "Imperfect Macroeconomic Expectations: Evidence and Theory", in Eichenbaum, M. and Hurst, E. (eds) *NBER Macroeconomics Annual* (Vol. 35, Chapter 1) (Chicago, IL: The University of Chicago Press) 1–86.

ANGELETOS, G.-M. and LIAN, C. (2018), "Forward Guidance Without Common Knowledge", *American Economic Review*, **108**, 2477–2512.

ANGELETOS, G.-M. and SASTRY, K. A. (2021), “Managing Expectations: Instruments vs. Targets”, *Quarterly Journal of Economics*, **136**, 2467–2532.

ANTOULAS, A. C. (2005), *Approximation of Large-Scale Dynamical Systems* (Philadelphia, PA: Society for Industrial and Applied Mathematics).

AOKI, M. (1983), “Equivalence of ARMA and State Space Models”, in Beckmann, M. and Krelle, W. (eds) *Notes on Economic Time Series Analysis: System Theoretic Perspectives* (Berlin: Springer-Verlag) 22–32.

BACCHETTA, P. and VAN WINCOOP, E. (2010), “Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle”, *American Economic Review*, **100**, 870–904.

BACCHETTA, P. and VAN WINCOOP, E. (2021), “Puzzling Exchange Rate Dynamics and Delayed Portfolio Adjustment”, *Journal of International Economics*, **131**, 103460.

BARBERIS, N., GREENWOOD, R., JIN, L., *et al.* (2015), “X-CAPM: An Extrapolative Capital Asset Pricing Model”, *Journal of Financial Economics*, **115**, 1–24.

BARBERIS, N., SHLEIFER, A. and VISHNY, R. (1998), “A Model of Investors Sentiment”, *Journal of Financial Economics*, **49**, 307–343.

BERK, R. H. (1966), “Limiting Behavior of Posterior Distributions When the Model is Incorrect”, *The Annals of Mathematical Statistics*, **37**, 51–58.

BORDALO, P., GENNAIOLI, N., MA, Y., *et al.* (2020), “Over-reaction in Macroeconomic Expectations”, *American Economic Review*, **110**, 2748–2782.

BORDALO, P., GENNAIOLI, N. and SHLEIFER, A. (2018), “Diagnostic Expectations and Credit Cycles”, *Journal of Finance*, **73**, 199–227.

BORDALO, P., GENNAIOLI, N. and SHLEIFER, A. (2020), “Memory, Attention, and Choice”, *Quarterly Journal of Economics*, **135**, 1399–1442.

CANDIAN, G. and LEO, P. DE (2023), “Imperfect Exchange Rate Expectations” (Working Paper, University of Maryland).

CHINN, M. D. and FRANKEL, J. A. (2020), “A Third of a Century of Currency Expectations Data: The Carry Trade and the Risk Premium” (Working Paper No. 2019-001, La Follette School).

COIBION, O. and GORODNICHENKO, Y. (2015), “Information Rigidity and the Expectations Formation Process: A Simple Framework and new Facts”, *American Economic Review*, **105**, 2644–2678.

Consensus Economics Inc. (2022), “Consensus Forecasts—G7 & Western Europe [dataset]” http://www.consensuseconomics.com/publications/consensus-forecasts-g7-western-europe/.

DA SILVEIRA, R. A., SUNG, Y. and WOODFORD, M. (2020), “Optimally Imprecise Memory and Biased Forecasts” (Working Paper No. 28075, NBER).

DONALD, S. G., FORTUNA, N. and PIPIRAS, V. (2007), “On Rank Estimation in Symmetric Matrices: The Case of Indefinite Matrix Estimators”, *Econometric Theory*, **23**, 1217–1232.

ENGEL, C. (2016), “Exchange Rates, Interest Rates, and the Risk Premium”, *American Economic Review*, **106**, 436–474.

ESPONDA, I. and POUZO, D. (2016), “Berk–Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models”, *Econometrica*, **84**, 1093–1130.

FAMA, E. F. (1984), “Forward and Spot Exchange Rates”, *Journal of Monetary Economics*, **14**, 319–338.

FARHI, E. and WERNING, I. (2019), “Monetary Policy, Bounded Rationality, and Incomplete Markets”, *American Economic Review*, **109**, 3887–3928.

FARMER, L. E., NAKAMURA, E. and STEINSSON, J. (2023), “Learning About the Long Run (Working Paper 29495, NBER).

FROOT, K. A. and FRANKEL, J. A. (1989), “Forward Discount Bias: Is it an Exchange Risk Premium?”, *Quarterly Journal of Economics*, **104**, 139–161.

FUSTER, A., HÉBERT, B. and LAIBSON, D. (2012), “Natural Expectations, Macroeconomic Dynamics, and Asset Pricing”, in Acemoglu, D. and Woodford, M. (eds) *NBER Macroeconomics Annual* (Vol. 26, Chapter 1) (Chicago, IL: The University of Chicago Press) 1–48.

FUSTER, A., LAIBSON, D. and MENDEL, B. (2010), “Natural Expectations and Macroeconomic Fluctuations”, *Journal of Economic Perspectives*, **24**, 67–84.

GABAIX, X. (2014), “A Sparsity-Based Model of Bounded Rationality”, *Quarterly Journal of Economics*, **129**, 1661–1710.

GOURINCHAS, P.-O. and TORNELL, A. (2004), “Exchange Rate Puzzles and Distorted Beliefs”, *Journal of International Economics*, **64**, 303–333.

GRANZIERA, E. and SIHVONEN, M. (2022), “Bonds, Currencies and Expectational Errors” (Research Paper 2020/3, Norges Bank).

HESPANHA, J. A. P. (2018), *Linear Systems Theory* (Princeton, NJ: Princeton University Press).

HIRSHLEIFER, D., LI, J. and YU, J. (2015), “Asset Pricing in Production Economies with Extrapolative Expectations”, *Journal of Monetary Economics*, **76**, 87–106.

HORN, R. A. and JOHNSON, C. R. (2013), *Matrix Analysis* (2nd edn) (New York, NY: Cambridge University Press).

KALEMLI-ÖZCAN, S. and VARELA, L. (2023), “Five Facts About the UIP Premium” (Working Paper No. 28923, NBER).

KOHLHAS, A. N. and WALTHER, A. (2021), “Asymmetric Attention”, *American Economic Review*, **111**, 2879–2925.

MOLAVI, P. (2022), “Simple Models and Biased Forecasts” (Working Paper, Northwestern University).

NAGEL, S. and XU, Z. (2022), “Asset Pricing with Fading Memory”, *Review of Financial Studies*, **35**, 2190–2245.

NEWEY, W. K. and WEST, K. D. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, **55**, 703–708.

PANAGEAS, S. (2020), "The Implications of Heterogeneity and Inequality for Asset Pricing", *Foundations and Trends in Finance*, **12**, 199–275.

PIAZZESI, M., SALOMAO, J. and SCHNEIDER, M. (2015), "Trend and Cycle in Bond Premia" (Working Paper, Stanford University).

RABIN, M. and VAYANOS, D. (2010), "The Gambler's and Hot-Hand Fallacies: Theory and Applications", *Review of Economic Studies*, **77**, 730–778.

SANDBERG, H. (2019), "Introduction to Model Order Reduction. Lecture Notes" http://www.diva-portal.org/smash/get/diva2:1338145/fulltext01.pdf.

SHALIZI, C. R. (2009), "Dynamics of Bayesian Updating with Dependent Data and Misspecified Models", *Electronic Journal of Statistics*, **3**, 1039–1074.

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

STAVRAKEVA, V. and TANG, J. (2020), "A Fundamental Connection: Survey-Based Exchange Rate Decomposition (Working Paper, London Business School).

VALCHEV, R. (2020), "Bond Convenience Yields and Exchange Rate Dynamics", *American Economic Journal: Macroeconomics*, **12**, 124–166.

WACHTER, J. A. and KAHANA, M. J. (2023), "A Retrieved-Context Theory of Financial Decisions" (Working Paper No. 26200, NBER).

WHITE, H. (1994), *Estimation, Inference and Specification Analysis*. Econometric Society Monographs No. 22. (Cambridge, UK: Cambridge University Press).