

# Model Complexity, Expectations, and Asset Prices\*

Pooya Molavi<sup>†</sup>

Alireza Tahbaz-Salehi<sup>‡</sup>

Andrea Vedolin<sup>§</sup>

November 2022

## Abstract

This paper analyzes 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.

---

\*We are grateful to three anonymous referees for helpful comments and suggestions. We also thank Fernando Alvarez, Turan Bali, Pedro Bordo, 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.

<sup>†</sup>Northwestern University, [pmolavi@kellogg.northwestern.edu](mailto:pmolavi@kellogg.northwestern.edu).

<sup>‡</sup>Northwestern University and CEPR, [alirezat@kellogg.northwestern.edu](mailto:alirezat@kellogg.northwestern.edu).

<sup>§</sup>Boston University, NBER, and CEPR, [avedolin@bu.edu](mailto:avedolin@bu.edu).

# 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.

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.<sup>1</sup>

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 behavioral 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

---

<sup>1</sup>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 ([Nagel and Xu, 2022](#); [da Silveira, Sung, and Woodford, 2020](#)), 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](#)).

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 behavioral 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  $k < n$ .

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 autocorrelation function 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 behavioral 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 behavioral 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.

**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 behavioral biases. For instance, [Barberis, Shleifer, and Vishny \(1998\)](#) 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.<sup>2</sup> 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, Li, and Yu \(2015\)](#), and [Barberis, Greenwood, Jin, and Shleifer \(2015\)](#) and diagnostic expectations models of [Bordalo, Gennaioli, and Shleifer \(2018\)](#) 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).

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 ([Nagel and Xu, 2022](#); [Bordalo, Gennaioli, and Shleifer, 2020](#); [Wachter and Kahana, 2021](#)), 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 behavioral 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

---

<sup>2</sup>For theoretical and empirical studies of departures from the full information rational expectations benchmark in the macroeconomics literature, see [Coibion and Gorodnichenko \(2015\)](#), [Bordalo, Gennaioli, Ma, and Shleifer \(2020\)](#), [Angeletos, Huo, and Sastry \(2020\)](#), [Kohlhas and Walther \(2021\)](#), and [Farmer, Nakamura, and Steinsson \(2022\)](#).

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.<sup>3</sup> 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.<sup>4</sup>

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 De Leo \(2022\)](#), and [Granziera and Sihvonen \(2022\)](#), who study how distortions in agents' subjective expectations can explain various exchange rate puzzles.<sup>5</sup> 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.

**Outline** The rest of the paper is organized as follows. Section 2 presents the environment and specifies our behavioral 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 behavioral 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 uncovered interest rate parity in foreign exchange markets. All proofs and some additional results are provided in the Appendix. An Online Appendix contains additional empirical results, as well as the proofs not presented in the main paper.

---

<sup>3</sup>We formalize this equivalence in Appendix 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.

<sup>4</sup>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.

<sup>5</sup>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.

## 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  $\{x_t\}_{t=-\infty}^{\infty}$  of variables, 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

$$\begin{aligned} z_t &= \mathbf{A}^* z_{t-1} + \mathbf{B}^* \epsilon_t \\ x_t &= c^{*'} z_t, \end{aligned} \tag{1}$$

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  $\epsilon_t \in \mathbb{R}^n$  are independent over time and across factors and are normally distributed with mean zero and unit variance, i.e.,  $\epsilon_{it} \sim \mathcal{N}(0, 1)$ . 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 entire family of stationary ARMA processes (Aoki, 1983).

Throughout, we assume that there are no redundant factors in (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 dividend at that time and its expected future price, discounted at some rate  $\delta$ . 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:

$$rx_{t+1} = \delta y_{t+1} - y_t + x_t. \quad (3)$$

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

Together with the specification of how the 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  $(z_1, \dots, z_n)$  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 behavioral 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 form

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

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. 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  $(\omega_1, \dots, \omega_k)$  in the agents' model may overlap with a subset of the  $n$  factors  $(z_1, \dots, z_n)$  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 favor of those that are more so (in a sense that we will formalize). Third, when  $k < n$ , the set of models entertained by agents does not contain the true  $n$ -factor data-generating process in (1), i.e.,  $\theta^* \notin \Theta_k$ . In such a case, our behavioral 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 behavioral 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  $\{x_{t_0}, \dots, x_{t-1}, x_t\}$ .

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 (**A**), volatility of and covariance between the noise terms (**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:

$$rx_{t+1} = \sum_{\tau=1}^{\infty} \delta^\tau (\mathbb{E}_{t+1}[x_{t+\tau}] - \mathbb{E}_t[x_{t+\tau}]). \quad (5)$$

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  $\{x_{t_0}, \dots, x_{t-1}, x_t\}$ , and consider the limit as  $t_0 \rightarrow -\infty$ . 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 the temporary fluctuations in beliefs that arise after observing a finite sample.

## 2.4 Discussion

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

As already mentioned, the behavioral 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.<sup>6</sup>

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.<sup>7</sup>

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

---

<sup>6</sup>In view of equations (2) and (3), this assumption also implies 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.

<sup>7</sup>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 nonzero posterior belief to models with more than  $k$  factors.

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 Appendix C.2, agents' forecasts may exhibit systematic over- or under-reaction to news depending on the statistical properties of the 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 (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 behavioral 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:

$$r_{x_{t+h}} = \alpha_h^{\text{rx}} + \beta_h^{\text{rx}} x_t + \epsilon_{t,h}, \quad (6)$$

where  $h \geq 1$  denotes the horizon and  $r_x$  is the excess return, as defined in (3). If returns are unpredictable—as would be the case under rational expectations—then  $\beta_h^{\text{rx}} = 0$  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  $\beta_h^{\text{rx}} > 0$ . Thus, the family of slope coefficients  $(\beta_1^{\text{rx}}, \beta_2^{\text{rx}}, \dots)$  not only provides a measure for departures from the rational expectations benchmark, but also determines the extent to which such departures vary with horizon  $h$ .<sup>8</sup>

---

<sup>8</sup>The family of regressions in (6) resembles Fama's (1984) regression specification for testing deviations from the uncovered interest rate parity 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  $r_{x_t}$  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.

### 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  $h \geq 1$ :

$$x_{t+h+m-1} - \mathbb{E}_{t+h-1}[x_{t+h+m-1}] = \alpha_{h,m}^{\text{fe}} + \beta_{h,m}^{\text{fe}} x_t + \epsilon_{t,h,m}, \quad (7)$$

where  $\mathbb{E}[\cdot]$  denotes the agents' subjective expectations. The unpredictability of forecast errors under rational expectations requires that  $\beta_{h,m}^{\text{fe}} = 0$  for all  $h, m \geq 1$ . Therefore, once again, the term structure of coefficients  $\beta_{h,m}^{\text{fe}}$  provides us with a natural measure for assessing the implications of our behavioral 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  $\tau$  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  $\xi_\tau^*$ . 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*

$$\beta_h^{\text{rx}} = \delta \frac{(\xi_h^* - \xi_h) - \sum_{\tau,s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_s (\xi_{h-\tau}^* - \xi_{h-\tau})}{1 - \sum_{\tau,s=1}^{\infty} \delta^\tau \Xi_{\tau s}^{-1} \xi_s} \quad (8)$$

and

$$\beta_{h,m}^{\text{fe}} = (\xi_{h+m-1}^* - \xi_{h+m-1}) - \sum_{\tau=1}^{\infty} \sum_{s=1}^{\infty} \xi_{s+m-1} \Xi_{\tau s}^{-1} (\xi_{h-\tau}^* - \xi_{h-\tau}), \quad (9)$$

respectively, where  $\Xi$  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  $\tau$  and hence,  $\beta_h^{\text{rx}} = \beta_{h,m}^{\text{fe}} = 0$  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-extrapolate the true persistence 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 the 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,  $\Theta_k$ . Given an arbitrary model  $\theta \in \Theta_k$  for the agents, denote the Kullback–Leibler (KL) divergence of model  $\theta$  from the true process by

$$\text{KL}(\theta^* \parallel \theta) = \mathbb{E}^*[-\log f^\theta(x_{t+1}|x_t, \dots)] - \mathbb{E}^*[-\log f^*(x_{t+1}|x_t, \dots)], \quad (10)$$

where  $f^*$  is the density of the fundamental under the true data-generating process,  $f^\theta$  is the agents' subjective density under model  $\theta$ , 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  $f^\theta$  and  $f^*$  coincide almost everywhere. As such,  $\text{KL}(\theta^* \parallel \theta)$  can be interpreted

as a measure for the disparity between agents' subjective expectations under model  $\theta$  and the true process.<sup>9</sup>

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

$$\text{ess inf}_{\theta \in \tilde{\Theta}} \text{KL}(\theta^* \parallel \theta) > \min_{\theta \in \hat{\Theta}_k} \text{KL}(\theta^* \parallel \theta),$$

then

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

This result, which is in line with the literature on learning under model misspecification (such as Berk (1966) and 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  $\theta^*$ . 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 behavioral 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

---

<sup>9</sup>To see this, note that  $\mathbb{E}^*[-\log f^\theta(x_{t+1}|x_t, \dots)]$  and  $\mathbb{E}^*[-\log f^*(x_{t+1}|x_t, \dots)]$  are entropy-like terms that capture the extent of uncertainty regarding one-step-ahead predictions conditional on the history of past observations under, respectively, model  $\theta$  and the true model,  $\theta^*$ . Therefore, the right-hand side of (10) measures the additional uncertainty—and hence, the resulting degradation in prediction quality—when agents use model  $\theta$  as opposed to the true model.

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.*

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

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

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

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

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

(b) *Furthermore, the coefficients of the return and forecast-error predictability regressions are given by*

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

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

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 (12) is in terms of matrix  $\mathbf{M}$  and vector  $u$ , these objects are closely linked to the parameters of agents' model,  $\theta = (\mathbf{A}, \mathbf{B}, c)$ .<sup>10</sup>

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 (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.

<sup>10</sup>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 (10).

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.*

(a) *If  $k \geq n$ , then  $\beta_h^{\text{rx}} = \beta_{h,m}^{\text{fe}} = 0$  for all horizons  $h$  and all  $m$ .*

(b) *If  $k < n$ , then there exist  $h, \tilde{h}$ , and  $m$  such that  $\beta_h^{\text{rx}} \neq 0$  and  $\beta_{h,m}^{\text{fe}} \neq 0$ .*

Statement (a) 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 the assumption that  $k \geq n$ , implies that agents rule out models that are inconsistent with their past observations in favor 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  $k > n$ ). Statement (b) 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  $\beta_h^{\text{rx}} = \beta_{h,m}^{\text{fe}} = 0$  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  $\tau$ ). 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:*

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

This result is a direct consequence of the assumptions that (i) the true data-generating process is stationary ergodic and (ii) agents only entertain stationary ergodic factor models. These assumptions imply that 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.

Next, we study how changes in the number of factors in agents' model determine forecast-error predictability. Let  $(\phi_1^*, \phi_2^*, \dots)$  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  $\varepsilon_t$ .<sup>11</sup> 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.

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



**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, (i)  $\bar{\beta}^{\text{fe}} > 0$  for all  $k < n$ ; (ii)  $\bar{\beta}^{\text{fe}}$  is decreasing in  $k$ ; and (iii)  $\bar{\beta}^{\text{fe}} = 0$  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

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

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

respectively, where  $\xi_h^*$  is the autocorrelation of the fundamental at lag  $h$ .

The closed-form 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, i.e., 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 (1), which governs the evolution of the fundamental, plays a central role in shaping the signs of the coefficients of the return predictability regression.

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

- (a) If both eigenvalues are real and  $\lambda_1 > 0$ , then  $\beta_h^{\text{rx}}$  never changes sign.
- (b) If both eigenvalues are real and  $\lambda_1 < 0$ , then  $\beta_h^{\text{rx}}$  changes sign at every horizon.
- (c) If both eigenvalues are complex, then  $\beta_h^{\text{rx}}$  changes sign at horizons  $h = r\pi/|\varphi|$  for  $r \in \mathbb{N}$ , where  $\varphi = \arg(\lambda_1)$  is the argument (or the phase) of  $\lambda_1$  in its polar representation.

That the return predictability pattern depends on the eigenvalues of  $\mathbf{A}^*$  is a consequence of the well-known fact that these eigenvalues are what determine the behavior of the autocorrelation of the data-generating process at different horizons. For example, when  $\mathbf{A}^*$  has a complex conjugate pair of eigenvalues, then  $\xi_h^*$  exhibits an oscillatory behavior, whereas it decays monotonically if both eigenvalues are real and positive. These patterns are then naturally reflected in the behavior of  $\beta_h^{\text{FX}}$ , as agents who are restricted to one-factor models fit a process with an exponentially decaying autocorrelation function to a process with a potentially 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, as represented by the maximum number of factors in their models. 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:

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

where  $(\phi_1^{(m)}, \phi_2^{(m)}, \dots)$  are a collection of coefficients that are independent of the fundamental's realizations. Given these coefficients, let  $\Phi$  denote the (infinite-dimensional) Hankel matrix with typical element  $\Phi_{ij} = \phi_{i+j-1}^{(m)}$  for all  $i, j \geq 1$ .<sup>12</sup> Note that, given the right- and the left-hand sides of (17), this matrix depends on the statistical properties of both the true data-generating process and the agents' subjective model. We have the following result.

**Proposition 8.** *Let  $k$  and  $n$  denote the number of factors in, respectively, agents' model and the true data-generating process. Then,*

$$\text{rank}(\Phi) = \min\{k, n\} \quad (18)$$

*generically, where  $\Phi$  is the Hankel matrix constructed from coefficients  $(\phi_1^{(m)}, \phi_2^{(m)}, \dots)$  in (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  $\Phi$ . The statement that  $\text{rank}(\Phi) = k$  means that all rows of  $\Phi$  can be expressed as linear combinations of its first  $k$  rows, which in turn implies that coefficients  $(\phi_{2k+1}^{(m)}, \phi_{2k+2}^{(m)}, \dots)$  cannot be determined independently from  $(\phi_1^{(m)}, \dots, \phi_{2k}^{(m)})$ .<sup>13</sup> Thus, an increase in  $\text{rank}(\Phi)$  translates into expanding the possibilities of how agents can make forecasts using their past observations in the form of (17)—as one would anticipate to be the case when agents rely on a richer set of models with a larger number of factors,  $k$ .

<sup>12</sup>We suppress the dependence of  $\Phi$  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  $j$ .

<sup>13</sup>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.

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  $k < n$ . When  $k \geq n$ , agents’ forecasts coincide with forecasts under rational expectations, which means agents form the same expectations irrespective of the actual value of  $k$ . This is why in the statement of Proposition 8,  $\text{rank}(\Phi) = \min\{k, n\}$  as opposed to  $k$ .<sup>14</sup> Second, note that when  $k < n$ , it may still be the case that the model in  $\Theta_k$  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 (10) generically falls in  $\Theta_k \setminus \Theta_{k-1}$ .

## 5 Application: Violations of Uncovered Interest Rate Parity

In this section, we apply our framework to study violations of uncovered interest rate parity (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 behavioral 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 four to seven 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 the high-interest rate country—cannot account for the predictability reversal puzzle.

To reproduce the empirical evidence on the above-mentioned UIP violations, we first map this

---

<sup>14</sup>In the special case that  $k = n$ , 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, e.g., [Antoulas \(2005, Theorem 4.40\)](#). Proposition 8 generalizes and adapts this result to our behavioral framework by allowing for expectations that are formed using a misspecified, less complex model.

context to our reduced-form environment in Subsection 2.1. Let the fundamental denote the log interest rate differential between the U.S. and a foreign country, i.e.,  $x_t = i_t^* - i_t$ , where  $i_t$  and  $i_t^*$  are nominal interest rates on deposits held in U.S. dollars and the foreign currency, respectively. Also, let  $y_t$  denote the log of the foreign exchange rate, expressed as the U.S. dollar 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:

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

According to this condition, an increase in the U.S. to foreign short-term interest rate differential is associated with an exchange rate appreciation, and a higher expected future exchange rate implies a depreciation. Note that, as in Subsection 2.1, the expectation in (19) denotes agents' 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:

$$rx_{t+1} = y_{t+1} - y_t + (i_t^* - i_t).$$

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 U.S. 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 from forward rates,  $i_t^* - i_t = f_t - y_t$ , also available from Datastream. We then run the following family of regressions:

$$rx_{t+h} = \alpha_h^{rx} + \beta_h^{rx}(i_t^* - i_t) + \epsilon_{t,h},$$

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 U.S.) lead to higher excess returns. This pattern remains the same up to a horizon of three years, but then reverses its sign, illustrating the predictability reversal puzzle: for horizons between four to seven years, higher interest rate differentials predict lower excess returns. Finally, as the figure indicates, the estimated coefficient  $\beta_h^{rx}$  becomes indistinguishable from zero at even longer horizons.

## 5.2 Return Predictability

Turning to our framework's predictions, we first calculate the autocorrelation function (ACF) of the interest rate differential between the trade-weighted average interest rate and the U.S. interest rate. Taking this autocorrelation,  $\{\zeta_\tau^*\}_{\tau \geq 1}$ , as our primitive, we then calculate the term structure of

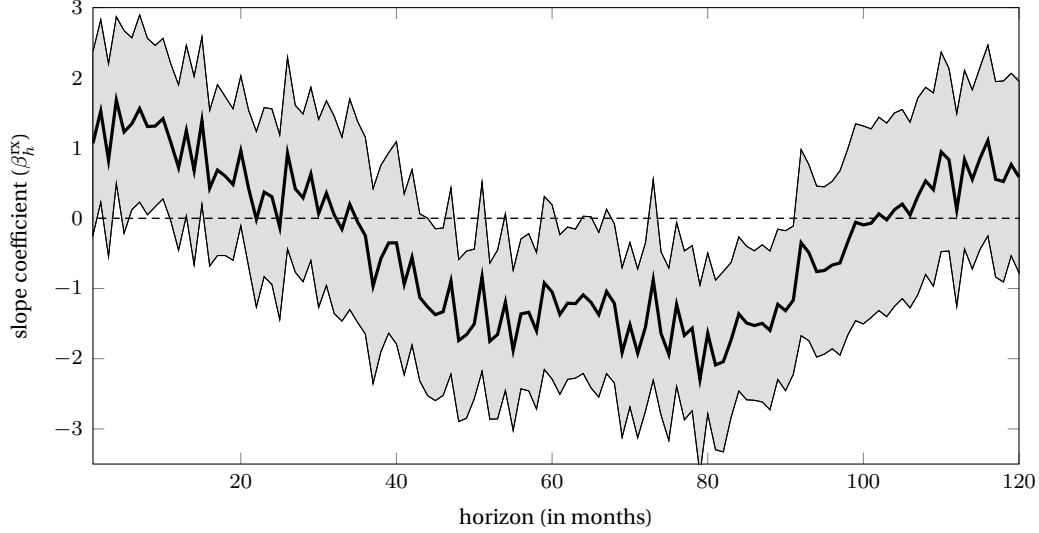


Figure 1. Violation of Uncovered Interest Rate Parity at Different Horizons

*Notes:* This figure plots estimated slope coefficients, together with 90 percent confidence intervals, of return predictability regression:  $rx_{t+h} = \alpha_h^{rx} + \beta_h^{rx}(i_t^* - i_t) + \epsilon_{t,h}$ , 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 U.S. dollar, and  $i_t^* - i_t$  is the log interest rate differential between the trade-weighted average interest rate and the U.S. interest rate. Monthly data from January 1985 to December 2021. Confidence intervals are calculated using [Newey and West \(1987\)](#) standard errors.

the model-implied coefficients  $\beta_h^{rx}$  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, i.e.,  $k = 1$ . 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 its magnitude).<sup>15</sup> 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.<sup>16</sup>

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).

<sup>15</sup>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 (2) abstracts from (subjective) risk premia. This assumption imposes an upper bound of  $|\beta_h^{rx}| \leq 1$  on model-implied slope coefficients in (15).

<sup>16</sup>In Appendix E, we redo this exercise separately for each of the currencies in our sample.

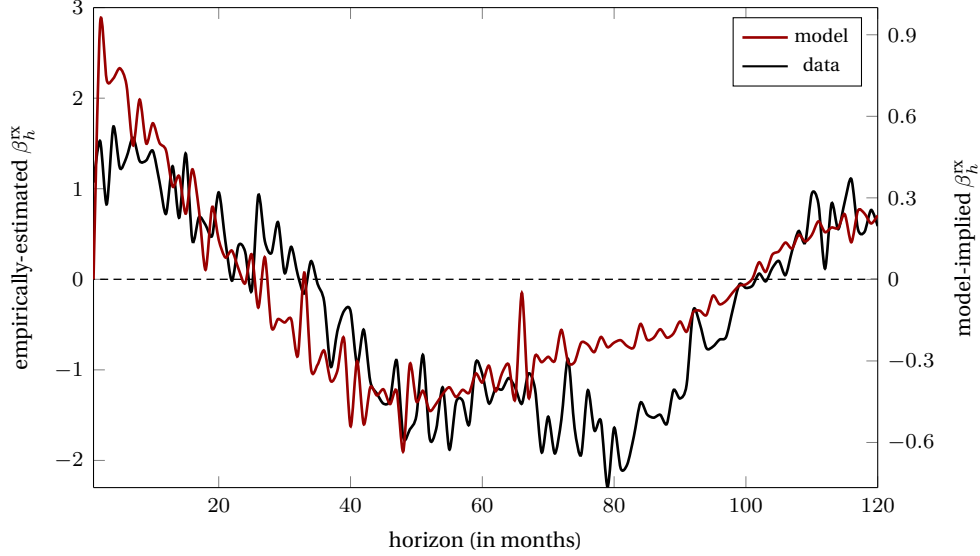


Figure 2. Term Structure of Return Predictability

*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.

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 autocorrelation function of the interest rate differential as an input to calculate  $\beta_h^{\text{rx}}$  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 equation (13) and the characterization result in Theorem 1 to solve for  $\beta_h^{\text{rx}}$  numerically.<sup>17</sup> As a by-product, we also obtain the model-implied autocorrelation function from the perspective of agents who are constrained to the class of  $k$ -factor models (equation (11)).

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 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 tracks the empirical ACF very closely, 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.

We can further explore the relationship between the interest rate differentials and the model-

<sup>17</sup>Note that the expressions in (11)–(13) involve infinite sums. To implement these expressions numerically, we truncate the infinite sums after 120 terms.

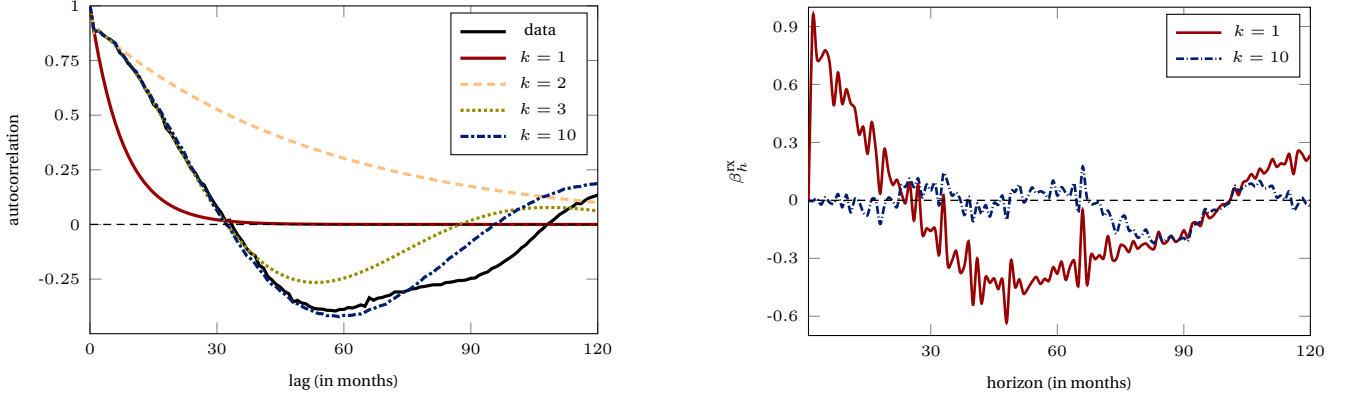


Figure 3. Varying the Complexity of Agents' Model

*Notes:* The left panel plots the autocorrelation function 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.

implied slope coefficients by using the result in Proposition 7, according to which the sign pattern of model-implied coefficients for  $k = 1$  is tightly linked to the spectrum of matrix  $A^*$  that governs the dynamics of the fundamental (in this case, the interest rate differential process). Given the close similarity between the ACFs of the 10-factor model and the interest rate differential process (Figure 3), we extract the eigenvalues corresponding to the estimated 10-factor model and treat them as approximations to the eigenvalues of  $A^*$ . We find that the two eigenvalues with the largest moduli form a complex conjugate pair with arguments  $\varphi = \pm 0.05$ . Therefore, to the extent that these two eigenvalues determine the dynamics of the true interest rate differential process, Proposition 7 implies that  $\beta_h^{\text{rx}}$  should change signs roughly every  $\pi/\varphi = 63$  months. This is consistent with the pattern in Figure 2, which exhibits two zero crossings within the 120-month horizon.

### 5.3 Forecast-Error Predictability

The findings in Subsection 5.2 indicate that our behavioral 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 behavioral 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 (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. Using these surveys, we calculate a “consensus forecast” for 3-month interest rates for Canada, the Eurozone, Japan, Switzerland, the

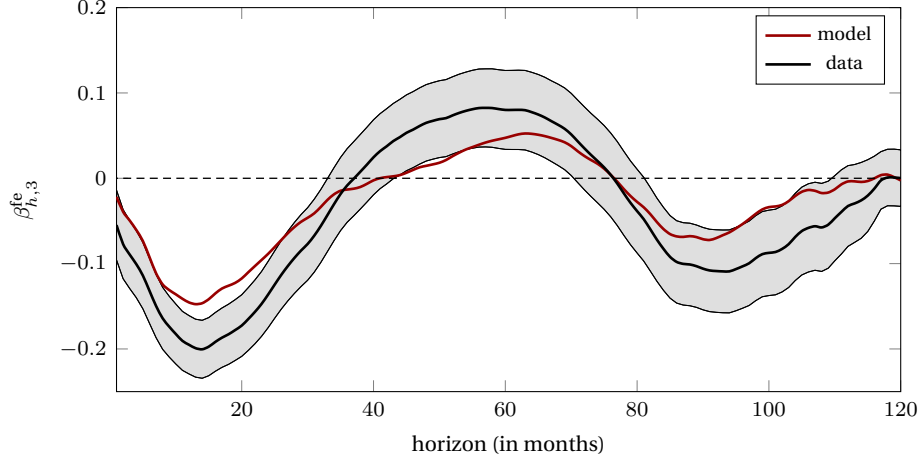


Figure 4. Term Structure of Forecast-Error Predictability

*Notes:* This figure plots the empirically-estimated (together with 95 percent 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.

United Kingdom, and the United States by taking the median of all forecasts for each country.<sup>18,19</sup> 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  $\beta_{h,3}^{fe}$  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 the two term structures in Figure 4 exhibit similar patterns, even though only the empirically-estimated coefficients are constructed from survey data: the model-implied coefficients in Figure 4 are simply obtained by 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  $\Phi$  constructed from the coefficients in (17) is simply equal to  $k$ . We

<sup>18</sup>Our dataset 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 Appendix D for more details about the dataset.

<sup>19</sup>Also see Stavrakeva and Tang (2020), Kalemli-Özcan and Varela (2021), and Candian and De Leo (2022) for other recent studies of determinants of exchange rate dynamics using survey data and Piazzesi, Salomao, and Schneider (2015) who use survey data to study of the relative importance of investor beliefs and risk assessment for bond return predictability.



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 from survey data on the past realizations of the interest rate differential to estimate coefficients  $(\phi_1^{(3)}, \phi_2^{(3)}, \dots)$  in (17); (ii) construct the Hankel matrix  $\Phi$  with typical element  $\Phi_{ij} = \phi_{i+j-1}^{(3)}$ ; and (iii) test for the rank of  $\Phi$ . 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  $\text{rank}(\Phi) = 1$ : the  $p$ -values corresponding to the rank test of Donald, Fortuna, and Pipiras (2007) never fall below 8%.<sup>20</sup> 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 behavioral biases.

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 uncovered interest rate parity 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.

---

<sup>20</sup>See Appendix E for the details. We also perform the same exercise for each country separately, finding similar results.

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 (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 behavioral 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 behavioral 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.

## A Appendix

### A.1 Model Complexity and Bounded Recall

The behavioral 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 appendix, 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 fundamentals:

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

where  $x_t$  is the realization of the fundamental at time  $t$ , and  $w = (w_1, \dots, w_k)$  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.<sup>21</sup> The key assumption is that the agent treats these moving averages as summary statistics for making predictions about all future realizations of

---

<sup>21</sup>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 vector  $w = (w_1, \dots, w_k)$  and matrix  $\mathbf{Q} = [q_{ij}]$  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  $t$ ).

the fundamental:

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

where  $\mathbb{E}_t[\cdot]$  denotes the agent's subjective expectation and the sequence of vectors  $v_\tau = (v_{\tau 1}, \dots, v_{\tau k})$  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 (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  $k \rightarrow \infty$  corresponding to an agent who can, in principle, recall the entire history of her observations.<sup>22</sup> Second, it is immediate to verify that the formulation in (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 setup, the agent represented by (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 (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 (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  $\Theta_k$  and  $\Psi_k$ .*

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).

Proposition A.1 has two important implications. First, it provides an alternative interpretation for the wedge between  $n$  and  $k$  and the restriction on the complexity of agents' models. In the main

---

<sup>22</sup>Note that the notion of bounded recall represented by (A.1) and (A.2) is different from the similarity-based recall model of Bordalo, Gennaioli, and Shleifer (2020) and the imprecise memory model of da Silveira, Sung, and Woodford (2020). Unlike these models, the behavioral 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.

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.

Second, Proposition A.1 implies that our results regarding the extent and nature of return predictability in Theorem 1 and Propositions 3–8 also apply to an agent who is restricted to making forecasts by relying on at most  $k$  statistics of the entire past history of her observations.

## 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 appendix, 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,  $\Theta_k$ , and let  $\hat{\Theta}_k$  denote an arbitrary compact subset of  $\Theta_k$ .*

(a) *Let  $\hat{\theta}_t^{\text{ML}} \in \arg \max_{\theta \in \hat{\Theta}_k} f^\theta(x_t, \dots, x_0)$  denote the agent's maximum likelihood estimator. Then,*

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

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

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

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.<sup>23</sup>

<sup>23</sup>Note that 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.

### 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 number of factors,  $k$ . In this appendix, we extend our previous results by assuming that only a fraction  $1 - \gamma$  of agents are subject to our behavioral constraint, while the remaining  $\gamma$  fraction can entertain models with any number of factors. For simplicity, we refer to the two groups of agents as behavioral and rational agents, respectively.<sup>24</sup>

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

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

where  $\bar{\mathbb{E}}[\cdot] = \gamma \mathbb{E}^*[\cdot] + (1 - \gamma) \mathbb{E}[\cdot]$  denotes the cross-sectional average of agents' expectations.<sup>25</sup> Iterating on the above, we can also obtain the following counterpart to equation (5) for excess returns in terms of agents' expectations:

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

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 [Barillas and Nimark \(2017\)](#), [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 behavioral agents, behavioral 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, behavioral agents are convinced—mistakenly so when  $k < n$ —that a  $k$ -factor model is sufficient to capture the process that drives the fundamental. We have the following result.

<sup>24</sup>Despite the terminology, recall that all agents in this economy are Bayesian, with the only difference between the two groups being that the “behavioral” 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 the same exact subjective expectations. The subjective expectations of the two groups differ only when  $k < n$ .

<sup>25</sup>We assume that  $\bar{\mathbb{E}}[\cdot]$  in equation (A.3) is the unweighted cross-sectional average of agents' expectations. See Appendix C.3 for a simple micro-foundation based on the overlapping generations model of [Allen, Morris, and Shin \(2006\)](#). 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.

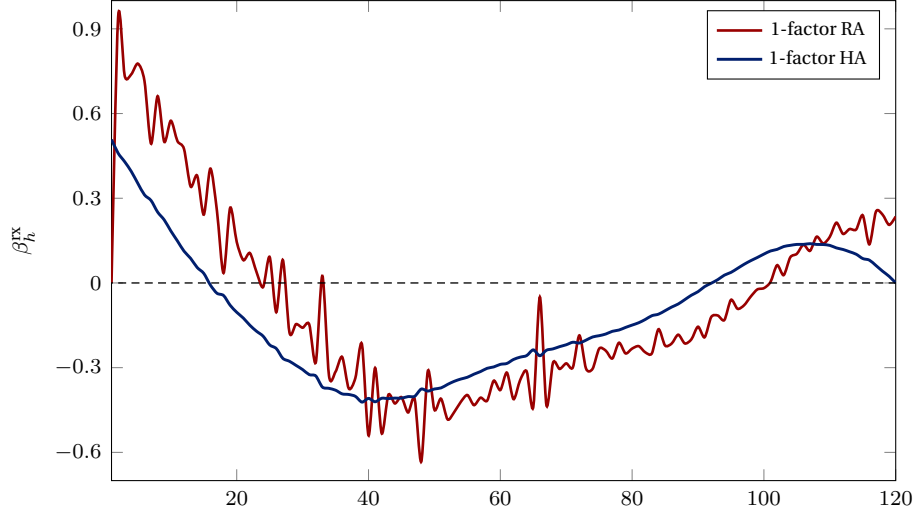


Figure A.1. Term Structures of Return Predictability in the Heterogeneous-Agent Economy

*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. Behavioral agents in both economies use a single-factor model. The fraction of behavioral agents in the heterogenous-agent economy is 10%, that is,  $\gamma = 0.9$ . Monthly data from January 1985 to December 2021.

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

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

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

The above result relates the term structure of  $\beta_h^{rx}$  in the heterogenous-agent economy to that of the economy only consisting of behavioral 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 behavioral agents. Note that, according to (A.5),  $\beta_h^{rx}(\gamma)$  is not simply a weighted average of  $\beta_h^{rx}(0)$  and  $\beta_h^{rx}(1) = 0$ . Rather, the fact that cross-sectional average expectations  $\mathbb{E}[\cdot]$  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 (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  $\gamma$  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% behavioral agents, i.e., 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 behavioral agents (i.e.,  $\gamma = 0$ ). This

indicates that even small fractions of behavioral agents can lead to notable deviations from the rational expectations benchmark.

## B Proofs and Derivations

### 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, one can express the agents'  $m$ -step-ahead forecast of the fundamental as follows:

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

where  $(\phi_1^{(m)}, \phi_2^{(m)}, \dots)$  is a sequence of constants 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

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

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  $\tau$  as perceived by the agents given their subjective model. Equation (B.2) thus relates the coefficients  $(\phi_1^{(m)}, \phi_2^{(m)}, \dots)$  in (B.1) to the subjective autocorrelations according to the model used by the agents.

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 (B.1) implies that

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

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

$$\beta_{h,m}^{\text{fe}} = (\xi_{h+m-1}^* - \xi_{h+m-1}) - \sum_{\tau=1}^{\infty} \phi_{\tau}^{(m)} (\xi_{h-\tau}^* - \xi_{h-\tau}) + \left( \xi_{h+m-1} - \sum_{\tau=1}^{\infty} \phi_{\tau}^{(m)} \xi_{h-\tau} \right), \quad (\text{B.3})$$

where  $\xi_{\tau}^* = \mathbb{E}^*[x_t x_{t+\tau}] / \mathbb{E}^*[x_t^2]$  is the autocorrelation of the fundamental at lag  $\tau$ . Note that, according to (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,

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

Next, note that, we can rewrite (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  $\Xi$  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 (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 (B.1) and for notational simplicity denoting  $\phi_r^{(1)}$  by  $\phi_r$ . Consequently,

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

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

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

Using (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,  $\text{rx}_{t+1} = \delta (x_{t+1} - \mathbb{E}_t[x_{t+1}]) + \text{rx}_{t+1} \sum_{r=1}^{\infty} \delta^r \phi_r$ , which in turn implies that  $\text{rx}_{t+1} = \delta (x_{t+1} - \mathbb{E}_t[x_{t+1}]) / (1 - \sum_{r=1}^{\infty} \delta^r \phi_r)$ . Hence, one can express the slope coefficients of the return predictability regression in terms of the slope coefficients of the forecast-error predictability regression as follows:

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

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

## 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

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

denote the log of the ratio of likelihoods of the fundamental under the true data-generating process ( $f^*$ ) and under the agents' subjective model ( $f^\theta$ ). 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,  $\lim_{t \rightarrow \infty} \frac{1}{t} \log f^\theta(x_t, x_{t-1}, \dots, x_0) = h(\theta^*, \theta)$  with  $\mathbb{P}^*$ -probability one uniformly in  $\theta \in \widehat{\Theta}_k$ , where  $h(\theta^*, \theta) = \mathbb{E}^*[\log f^\theta(x_{t+1}|x_t, x_{t-1}, \dots)]$ . This implies that

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

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

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

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

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

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 for any  $\epsilon$  and large enough  $t$ ,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \dots, x_0)}{f^*(x_t, \dots, x_0)} d\mu_0(\theta) \leq - \min_{\theta \in \widehat{\Theta}_k} \text{KL}(\theta^* || \theta)$$

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

$$\tau(G, \delta) = \max \left\{ t : \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \dots, x_0)}{f^*(x_t, \dots, x_0)} \mathbf{1}_{\{\theta \in G\}} d\mu_0(\theta) > \delta + \limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_{\widehat{\Theta}_k} \frac{f^\theta(x_t, \dots, x_0)}{f^*(x_t, \dots, x_0)} \mathbf{1}_{\{\theta \in G\}} d\mu_0(\theta) \right\}.$$

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.  $\square$

## Proof of Theorem 1

**Proof of part (a)** By Proposition 2, agents' long-run beliefs concentrate on the set of models with minimum KL divergence from the true model,  $\theta^*$ . We thus start by characterizing  $\widehat{\Theta}_k = \text{argmin}_{\theta \in \Theta_k} \text{KL}(\theta^* || \theta)$ , where the KL divergence between the two models is defined in (10). As a first observation, note that instead of optimizing over  $\Theta_k$ , we can optimize over  $\underline{\Theta}_1 \cup \underline{\Theta}_2 \cup \dots \cup \underline{\Theta}_k$ , 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.

Next, note that under model  $\theta$ , agents believe that the fundamental is described by the process in (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  $\omega_{t+1}$  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 semidefinite matrix that satisfies the algebraic Riccati equation

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

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

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

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

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

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

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

To minimize the KL divergence between agents' model and the true data-generating process over  $\Theta_r$ , it is convenient to normalize the model parameters via a change of variables. In particular, let

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

where  $\mathbf{R}$  is a positive definite matrix such that  $\mathbf{R} c = \hat{\Sigma} c$ .<sup>26</sup> Note that  $u$  is a  $r$ -dimensional vector of unit length and  $\mathbf{M}$  is a  $r \times r$  stable matrix. Given the change of variables in (B.11), agents'  $s$ -step ahead forecasts in (B.9) can be written as

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

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 the KL divergence between agents' model and the true data-generating process is given by

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

<sup>26</sup>Since  $c' \hat{\Sigma} c > 0$ , such a matrix  $\mathbf{R}$  is always guaranteed to exist. See Lemma C.2 for a proof.

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

$$\begin{aligned} \text{KL}(\theta^* \parallel \theta) &= \mathbb{E}^*[\log f^*(x_{t+1}|x_t, \dots)] + \log \sqrt{2\pi e} + \frac{1}{2} \log \mathbb{E}^*[x_t^2] \\ &+ \frac{1}{2} \log \left( 1 - 2 \sum_{s=1}^{\infty} \phi_s \xi_s^* + \sum_{\tau=1}^{\infty} \phi_s \phi_{\tau} \xi_{\tau-s}^* \right). \end{aligned} \quad (\text{B.14})$$

over  $(\mathbf{M}, u)$ , where  $\xi_s^*$  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 (B.14) depends on agents' model, it follows immediately that minimizing the KL divergence between a  $k$ -factor model  $\theta \in \Theta_k$  and the true underlying model  $\theta^*$  is equivalent to minimizing (12) over  $\mathbf{M}$  and  $u$ .

Having determined  $\mathbf{M}$  and  $u$  as the solutions to the optimization in (12), we next derive the expression in (11) for the autocorrelation of the fundamental according to agents' subjective model. Equation (4) implies that the autocovariance of the fundamental according under agents' subjective model is given by  $\mathbb{E}[x_t x_{t-h}] = c' \mathbb{E}[\omega_t \omega_{t-h}'] c = c' \mathbf{A}^h \mathbb{E}[\omega_t \omega_t'] c$  for all  $h \geq 0$ . Therefore, applying the change of variables in (B.11) implies that the autocorrelation of the subjective autocorrelation at lag  $h \geq 0$  is given by

$$\xi_h = \frac{u' \mathbf{M}^h \mathbf{G} u}{u' \mathbf{G} u}, \quad (\text{B.15})$$

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

$$\mathbf{G} = \sum_{s=0}^{\infty} \mathbf{M}^s (\mathbf{R}^{-1/2} \mathbf{B} \mathbf{B}' \mathbf{R}^{-1/2}) \mathbf{M}'^s. \quad (\text{B.16})$$

Next, note that multiplying both sides of the algebraic Riccati equation in (B.7) by  $\mathbf{R}^{-1/2}$  from the left and the right implies that

$$\mathbf{R}^{-1/2} \mathbf{B} \mathbf{B}' \mathbf{R}^{-1/2} = \mathbf{R}^{-1/2} \hat{\Sigma} \mathbf{R}^{-1/2} - \mathbf{M} \mathbf{R}^{-1/2} \hat{\Sigma} \mathbf{R}^{-1/2} \mathbf{M}' + \mathbf{M} u u' \mathbf{M}',$$

where we are using the expressions in (B.11) and the fact that  $\hat{\Sigma} c = \mathbf{R} c$ . Plugging the above into the expression for  $\mathbf{G}$  in (B.16) and multiplying both sides by  $u$  from the right, we obtain

$$\mathbf{G} u = \mathbf{R}^{-1/2} \hat{\Sigma} \mathbf{R}^{-1/2} u + \sum_{s=1}^{\infty} \mathbf{M}^s u u' \mathbf{M}'^s u.$$

Note that  $\hat{\Sigma} \mathbf{R}^{-1/2} u = \hat{\Sigma} c / \sqrt{c' \mathbf{R} c} = \mathbf{R} c / \sqrt{c' \mathbf{R} c} = \mathbf{R}^{1/2} u$ . Hence,  $\mathbf{G} u = u + \sum_{s=1}^{\infty} \mathbf{M}^s u u' \mathbf{M}'^s u$ . Plugging this expression into (B.15) and using the fact that  $u' u = 1$  then establishes (11).  $\square$

**Proof of part (b)** On the one hand, according to equation (B.1), agents'  $m$ -step-ahead forecast can be expressed as  $\mathbb{E}_{t-1}[x_{t+m-1}] = \sum_{\tau=1}^{\infty} \phi_{\tau}^{(m)} x_{t-\tau}$ . On the other hand, according to (B.12), these same forecasts can be expressed as  $\mathbb{E}_{t-1}[x_{t+m-1}] = \sum_{\tau=1}^{\infty} u' \mathbf{M}^{m-1} [\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M} u x_{t-\tau}$ . Matching the coefficients of these two equations therefore implies that  $\phi_{\tau}^{(m)} = u' \mathbf{M}^{m-1} [\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M} u$  for all  $\tau \geq 1$ .

Next, recall from the proof of Proposition 1 that the slope coefficient of the forecast-error predictability regression is  $\beta_{h,m}^{\text{fe}} = \xi_{h+m-1}^* - \sum_{\tau=1}^{\infty} \phi_{\tau}^{(m)} \xi_{h-\tau}^*$ . Using the fact that  $\phi_{\tau}^{(m)} = u' \mathbf{M}^{m-1} [\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M} u$  then establishes (14). To establish (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 (B.5). Replacing for  $\beta_{h,1}^{\text{fe}}$  from (14) and using the fact that  $\phi_{\tau} = u' [\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M} u$  then establishes (13).  $\square$

### Proof of Proposition 3

**Proof of part (a)** When  $k \geq n$ , the true model is within the set of models considered by the agents, i.e.,  $\theta^* \in \Theta_n \subseteq \Theta_k$ . As a result, agents' subjective expectations coincide with rational expectations, which guarantees that  $\xi_{\tau} = \xi_{\tau}^*$  for all  $\tau$ . Therefore, equations (8) and (9) in Proposition 1 implies that  $\beta_h^{\text{rx}} = \beta_{h,m}^{\text{fe}} = 0$  for all  $h \geq 1$ .  $\square$

**Proof of part (b)** We first show that if  $\beta_{h,m}^{\text{fe}} = 0$  for all  $h$  and  $m$ , then  $k \geq n$ . Suppose to the contrary that  $k < n$  and define  $\phi_s$  and  $\phi_s^*$  to denote the output from the minimization in (12) in Theorem 1 for the  $k$ - and  $n$ -factor models, respectively. By part (a) of the theorem, the slope coefficients of the forecast-error predictability regression arising from the optimal  $n$ -factor model are equal to zero at all horizons. Therefore, (14) implies that

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

Furthermore, by assumption,  $\beta_{h,1}^{\text{fe}} = 0$  for all  $h$  under the  $k$ -factor model. Therefore, (13) implies that

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

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

$$\sum_{h=1}^{\infty} \phi_h^* \xi_h^* = \sum_{h=1}^{\infty} \phi_h \xi_h^*. \quad (\text{B.19})$$

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

$$H^* = 1 - \sum_{s=1}^{\infty} \phi_s^* \xi_s^* \quad , \quad H = 1 - \sum_{s=1}^{\infty} \phi_s \xi_s^*.$$

Comparing the above two equations with (B.19) implies that  $H = H^*$ . Consequently, the  $k$ -factor model results in the same KL divergence from the data-generating process as the  $n$ -factor model, whose KL divergence from the data-generating process is equal to zero by assumption. This means that 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  $\beta_h^{\text{rx}} \neq 0$ .  $\square$

#### Proof of Proposition 4

Since the process that generates the fundamental is stationary and ergodic, its autocorrelation function  $\xi_s^* = \mathbb{E}^*[x_t x_{t+s}] / \mathbb{E}^*[x_t^2]$  decays at an exponential rate as  $s \rightarrow \infty$ . Next, we show that  $\phi_s = u'[\mathbf{M}(\mathbf{I} - uu')]^{s-1} \mathbf{M}u$  also decays at an exponential rate, where  $\mathbf{M}$  and  $u$  are given by (B.11). To this end, first note that,  $\mathbf{M}(\mathbf{I} - uu') = \mathbf{R}^{-1/2}(\mathbf{A} - g'c)\mathbf{R}^{1/2}$ . Since all eigenvalues of  $\mathbf{A} - g'c$  are inside the unit circle, it is immediate that  $\phi_s$  decays exponentially in  $s$ .

With the above in hand, we next show that  $\lim_{h \rightarrow \infty} \beta_h^{\text{rx}} = 0$ . Recall from Theorem 1 that the slope coefficient of the return predictability regression satisfies (13). Furthermore,  $\phi_\tau = u'[\mathbf{M}(\mathbf{I} - uu')]^{\tau-1} \mathbf{M}u$ , it follows that

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

where the second equality follows from an application of the Sherman-Morrison formula. Therefore, (13) implies that

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

and hence, by triangle inequality,

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

Since  $\xi_s^*$  and  $\phi_s$  converges to zero at exponential rates, there are constants  $c_1, c_2 > 0$  and  $\rho_1, \rho_2 < 1$  such that  $|\xi_s^*| \leq c_1 \rho_1^s$  and  $|\phi_s| \leq c_2 \rho_2^s$  for all  $s$ . Consequently,

$$|\beta_h^{\text{rx}}| \leq \delta |u'(\mathbf{I} - \delta \mathbf{M})^{-1} u| \left( c_1 \rho_1^h + c_1 c_2 (\max\{\rho_1, \rho_2\})^h \sum_{\tau=1}^h \left( \frac{\min\{\rho_1, \rho_2\}}{\max\{\rho_1, \rho_2\}} \right)^\tau + c_1 c_2 \rho_2^h \sum_{\tau=1}^{\infty} (\rho_1 \rho_2)^\tau \right),$$

and as a result,  $|\beta_h^{\text{rx}}| \leq c_3 h \rho_3^h$  for all  $h$  for some constant  $c_3 > 0$  and  $\rho_3 = \max\{\rho_1, \rho_2\}$ . This inequality then guarantees that  $\lim_{h \rightarrow \infty} \beta_h^{\text{rx}} = 0$ . A similar argument also implies that  $\lim_{h \rightarrow \infty} \beta_{h,m}^{\text{fe}} = 0$ .  $\square$

#### Proof of Proposition 5

We start by establishing the following lemma.

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

*Proof.* The first-order condition of (12) with respect to  $\mathbf{M}'$  implies that

$$\sum_{s=1}^{\infty} \frac{d\phi_s}{d\mathbf{M}'} \left( \xi_s^* - \sum_{\tau=1}^{\infty} \phi_{\tau} \xi_{s-\tau}^* \right) = 0,$$

which given equation (14) can be rewritten as

$$\sum_{s=1}^{\infty} \frac{d\phi_s}{d\mathbf{M}'} \beta_{s,1}^{\text{fe}} = 0.$$

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  $\sum_{s=1}^{\infty} \phi_s \beta_{s,1}^{\text{fe}} = 0$ .  $\square$

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

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

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,

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

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

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

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

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

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,

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

Finally, note that since  $(\phi_1^*, \phi_2^*, \dots)$  are the coefficients of the autoregressive representation of the true process,  $\mathbb{E}_{t-1}^*[x_t] = \sum_{h=1}^{\infty} \phi_h^* x_{t-h}$ , which implies that  $\xi_\tau^* = \sum_{h=1}^{\infty} \phi_h^* \xi_{h-\tau}^*$ . Using this equality to simplify the above equation therefore leads to

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

Hence,

$$\bar{\beta}^{\text{fe}} = \left( 1 - \sum_{\tau=1}^{\infty} \phi_\tau^* \xi_\tau^* \right) \left( e^{2\text{KL}(\theta^* \|\theta)} - 1 \right). \quad (\text{B.20})$$

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 (B.20) is independent of agents' model, it follows that minimizing the KL divergence between agents' model and the true model is equivalent to minimizing  $\bar{\beta}^{\text{fe}}$ . Therefore, Proposition 2 implies that increasing the number of factors,  $k$ , in agents' model can only result in a reduction in  $\bar{\beta}^{\text{fe}}$ .  $\square$

## Proof of Proposition 6

Recall from Theorem 1 that the slope coefficients of the return and forecast-error predictability regressions are given by (13) and (14), where the sequence  $(\phi_1, \phi_2, \dots)$  is obtained by minimizing (12) over the  $k \times k$  stable matrix  $\mathbf{M}$  and the unit vector  $u \in \mathbb{R}^k$ . Therefore, when  $k = 1$ , it must be the case that  $\mathbf{M}$  is a scalar, denoted by  $\mu$ , satisfying  $|\mu| < 1$  and  $u \in \{-1, 1\}$ . This immediately 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, the expressions for the slope coefficients of predictability regressions in (13) and (14) reduce to

$$\beta_h^{\text{rx}} = \frac{\delta}{1 - \delta\mu} (\xi_h^* - \mu \xi_{h-1}^*) \quad (\text{B.21})$$

$$\beta_{h,m}^{\text{fe}} = \xi_{h+m-1}^* - \mu^m \xi_{h-1}^*, \quad (\text{B.22})$$

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

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

Optimizing the above over  $\mu \in (-1, 1)$  implies that  $\mu = \xi_1^*$ . Plugging in the results into (B.21) and (B.22) then completes the proof.  $\square$

## 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, i.e.,  $\mathbf{A}^{*n} = \sum_{i=1}^n \gamma_i \mathbf{A}^{*(n-i)}$ . Multiplying both sides of this equation by  $\mathbf{A}^{*h}$  implies that

$$\mathbf{A}^{*h} = \gamma_1 \mathbf{A}^{*(h-1)} + \dots + \gamma_n \mathbf{A}^{*(h-n)} \quad (\text{B.24})$$

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

$$\xi_h^* = \gamma_1 \xi_{h-1}^* + \dots + \gamma_n \xi_{h-n}^* \quad (\text{B.25})$$

for all  $h \geq n$ . On the other hand, recall from (15) that when  $k = 1$ , the slope coefficient of the return predictability regression at horizon  $h$  is given by  $\beta_h^{\text{rx}} = \delta(\xi_h^* - \xi_1^* \xi_{h-1}^*) / (1 - \delta \xi_1^*)$ . This observation, combined with (B.25) then implies that

$$\beta_h^{\text{rx}} = \gamma_1 \beta_{h-1}^{\text{rx}} + \dots + \gamma_n \beta_{h-n}^{\text{rx}} \quad (\text{B.26})$$

for all  $h \geq n + 1$ . In other words, the slope coefficients of the return predictability regression satisfy 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 (B.26) when all eigenvalues of  $\mathbf{A}^*$  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 horizon  $h$ . Furthermore, note that, by (15),  $\beta_1^{\text{rx}} = 0$ , which implies that  $\sum_{i=1}^n \zeta_i = 0$ . Therefore, in the special case that  $n = 2$ , 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  $\zeta_1$  is some constant and  $\lambda_1$  and  $\lambda_2$  are the eigenvalues of  $\mathbf{A}^*$ . Consequently,

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

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

**Proof of part (a)** Suppose both  $\lambda_1$  and  $\lambda_2$  are real and that  $\lambda_1 > 0$ . It is immediate that the right-hand side of (B.27) never changes sign. Therefore,  $\beta_h^{\text{rx}}$  has the same sign as  $\beta_2^{\text{rx}}$  for all  $h$ .  $\square$



**Proof of part (b)** Suppose both  $\lambda_1$  and  $\lambda_2$  are real and that  $\lambda_1 < 0$ . In this case, the right-hand side of (B.27) has the same sign as  $\beta_2^{\text{rx}}$  when  $h$  is even, whereas it has the opposite sign of  $\beta_2^{\text{rx}}$  when  $h$  is odd. This means that  $\beta_h^{\text{rx}}$  changes sign at every horizon.  $\square$

**Proof of part (c)** Suppose both  $\lambda_1$  and  $\lambda_2$  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 (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,  $\beta_{h+1}^{\text{rx}}$  has the same sign as  $\beta_2^{\text{rx}}$  if  $\sin(h\varphi) > 0$  and has the opposite sign if  $\beta_2^{\text{rx}}$  if  $\sin(h\varphi) < 0$ . Therefore,  $\beta_{h+1}^{\text{rx}}$  changes sign at frequency  $\pi/\varphi$ .  $\square$

### Proof of Proposition 8

We start by showing that, given a generic  $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$ , agents' beliefs concentrate on a set of models with a minimal representation in  $\Theta_p$ , where  $p = \min\{k, n\}$ . To show this, we consider two separate cases. If  $k \geq n$ , the assumption that the true model has a minimal representation of order  $n$ , coupled with the fact that agents recover the true model (Proposition 2), implies that agents' belief concentrate on the set of models with a minimal representation in  $\Theta_n = \Theta_p$ . If  $k < n$ , then Proposition 2 implies that agents' long-run beliefs concentrate on the set of models in  $\Theta_k$  with minimum KL divergence from  $\theta^*$ . Crucially, for a generic  $\theta^*$ , the models  $\theta \in \Theta_k$  that minimize  $\text{KL}(\theta^* \parallel \theta)$  belong to  $\Theta_k \setminus \Theta_{k-1}$ . Therefore, taken together, these two cases imply that, given a generic  $\theta^*$ , agents' beliefs concentrate on the set of models with a minimal representation in  $\Theta_p$ . This means that, without loss of generality, we can represent the agents' model with a minimal representation  $\theta \in \Theta_p$ .

Let  $(\mathbf{M}, u)$  denote the  $p$ -dimensional representation of agents' model  $\theta$ , where  $p = \min\{k, n\}$ ,  $\mathbf{M}$  is a  $p \times p$  stable matrix,  $u \in \mathbb{R}^p$  is a vector of unit length, and the pair  $(\mathbf{M}, u)$  is the solution to the optimization problem in (12).<sup>27</sup> Given this representation, equation (B.12) implies 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 (17) implies that coefficients  $(\phi_1^{(m)}, \phi_2^{(m)}, \dots)$  can be expressed in terms of  $\mathbf{M}$  and  $u$  as

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

<sup>27</sup>Recall that the pair  $(\mathbf{M}, u)$  relates to the parameters of agents model  $\theta = (\mathbf{A}, \mathbf{B}, c) \in \Theta_p$  via the change of variables in (B.11).

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

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

$$\Phi^{(1)} = \mathbf{O}\mathbf{C}, \quad (\text{B.29})$$

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

$$\mathbf{O} = \begin{bmatrix} u' \\ u'\mathbf{S} \\ u'\mathbf{S}^2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \mathbf{C} = [\mathbf{M}u \quad \mathbf{S}\mathbf{M}u \quad \mathbf{S}^2\mathbf{M}u \quad \dots],$$

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

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

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

We first show that, generically,  $\text{rank}(\mathbf{O}) = p$ . Suppose to the contrary that  $\text{rank}(\mathbf{O}) < p$ , 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 nonzero vector  $w$  and a scalar  $\lambda$  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  $\text{rank}(\mathbf{O}) = p$ .

Next, we establish that, generically,  $\text{rank}(\mathbf{C}) = p$ . Suppose to the contrary that  $\text{rank}(\mathbf{C}) < p$ , 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 nonzero vector  $w$  and a scalar  $\lambda$  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  $\theta^* \in \Theta_n$ , it must be the case that  $\text{rank}(\mathbf{C}) = p$ . This, together with the fact that  $\text{rank}(\mathbf{O}) = p$ , establishes that  $\text{rank}(\Phi^{(1)}) = p$  generically.

We now show that  $\text{rank}(\Phi^{(m)}) = p$  for  $m \geq 2$ . Start with the observation that the expression in (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  $\phi_1^{(0)} = 1$ . Given that  $\Phi_{i,j}^{(m)} = \phi_{i+j-1}^{(m)}$ , we can write the elements of  $\Phi^{(m)}$  in terms of the elements of  $\Phi^{(1)}$  as follows:

$$\Phi_{i,j}^{(m)} = \sum_{r=0}^{m-1} \phi_1^{(m-r-1)} \Phi_{i+r,j}^{(1)}$$

for all  $m \geq 2$  and all  $i, j \geq 1$ . This means that  $\Phi^{(m)} = \mathbf{L}^{(m)} \Phi^{(1)}$  for an upper-triangular matrix  $\mathbf{L}^{(m)}$ , whose diagonal elements are all equal to  $\phi_1^{(m-1)}$ . Consequently,  $\text{rank}(\Phi^{(m)}) = \text{rank}(\Phi^{(1)}) = p$  as long as  $\phi_1^{(m-1)} \neq 0$ . Finally, note that  $\phi_1^{(m-1)} = u' \mathbf{M}^{m-1} u$ , which is different from zero generically. Thus, generically,  $\text{rank}(\Phi^{(m)}) = p$  for all  $m \geq 2$ .  $\square$

### 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, i.e.,  $e_1' \mathbf{V}(\mathbf{A} - gc') = 0$ , 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:

$$s_t = \mathbf{P}g x_t + \mathbf{P}(\mathbf{A} - gc') \mathbf{P}^{-1} s_{t-1}, \quad (\text{B.30})$$

which can be rewritten as  $s_t = \sum_{r=0}^{\infty} [\mathbf{P}(\mathbf{A} - gc') \mathbf{P}^{-1}]^r \mathbf{P}g x_{t-r} = \mathbf{P} \sum_{r=0}^{\infty} (\mathbf{A} - gc')^r g x_{t-r}$ . On the other hand, 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). As a result, one can write the agent's forecasts in terms of vector  $s_t$ :

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

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  $s_{t+1} = w x_{t+1} + \mathbf{Q} s_t$ , 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

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

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:

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

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 (B.32) can be rewritten as

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

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  $s_t = \mathbf{V} \sum_{r=0}^{\infty} (\mathbf{A} - gc')^r g x_{t-r}$ . Multiplying both sides of this equation by  $(\mathbf{I} - e_1e_1')$  from the left implies that

$$(\mathbf{I} - e_1e_1')s_t = \mathbf{V} \sum_{r=0}^{\infty} (\mathbf{A} - gc')^r g x_{t-r} - e_1e_1' \mathbf{V}g x_t = \mathbf{V} \sum_{r=0}^{\infty} (\mathbf{A} - gc')^r g x_{t-r}, \quad (\text{B.33})$$

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  $e_1' \mathbf{V}g = 0$ . On the other hand, 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  $\sum_{r=0}^{\infty} (\mathbf{A} - gc')^r g x_{t-r}$  with  $\mathbf{V}^{-1}(\mathbf{I} - e_1e_1')s_t$  from (B.33) into (B.9) implies that the agent's forecasts can be written in terms of the vector of moving averages  $s_t$  in (B.32):

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

Hence,

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

Comparing equations (B.32)–(B.34) to equations (A.1)–(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 (A.1) and makes forecasts according to (A.2). Rewriting (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  $v_\tau$  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  $v_\tau'$  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,  $v'_{\tau+1} = (v'_\tau w) v'_1 + v'_\tau \mathbf{Q}$  for all  $\tau \geq 1$ . Therefore,  $v'_{\tau+1} = v'_\tau (wv'_1 + \mathbf{Q})$ , 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

$$\mathbb{E}_t[x_{t+\tau}] = v'_1 (wv'_1 + \mathbf{Q})^{\tau-1} s_t = v'_1 (wv'_1 + \mathbf{Q})^{\tau-1} \sum_{r=0}^{\infty} \mathbf{Q}^r w x_{t-r} \quad (\text{B.35})$$

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 (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

$$\mathbf{A} = wv'_1 + \mathbf{Q}, \quad \mathbf{B} = pe'_1, \quad c' = v'_1 \quad (\text{B.36})$$

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 (B.36), it follows from (B.8) that the corresponding Kalman gain is given by

$$g = (wv'_1 + \mathbf{Q}) \hat{\Sigma} v_1 (v'_1 \hat{\Sigma} v_1)^{-1}, \quad (\text{B.37})$$

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

$$\begin{aligned} \hat{\Sigma} &= (wv'_1 + \mathbf{Q}) \left( \hat{\Sigma} - \frac{1}{v'_1 \hat{\Sigma} v_1} \hat{\Sigma} v_1 v'_1 \hat{\Sigma} \right) (v_1 w' + \mathbf{Q}') + pp' \\ &= \mathbf{Q} \left( \hat{\Sigma} - \frac{1}{v'_1 \hat{\Sigma} v_1} \hat{\Sigma} v_1 v'_1 \hat{\Sigma} \right) \mathbf{Q}' + pp'. \end{aligned}$$

Recall that, by assumption,  $\mathbf{Q}p = 0$ . It is therefore immediate to verify that the positive semidefinite matrix  $\hat{\Sigma} = pp'$  satisfies the above equation. Therefore, the Kalman gain in (B.37) is given by  $g = w$ . Replacing for  $g$  and the parameters in (B.36) into the expression for the agent's forecasts in (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 (B.35). In other words, the forecasts of the agent with the  $k$ -factor model with parameters in (B.36) coincide with those of the agent with  $k$  moving averages with parameters given by  $(w, \mathbf{Q}, \{v_\tau\})$ .  $\square$

## Proof of Proposition A.2

**Proof of part (a)** Fix an arbitrary compact subset  $\hat{\Theta}_k \subseteq \Theta_k$  of the class of  $k$ -factor models and let  $\hat{\theta}_t^{\text{ML}} \in \arg \max_{\theta \in \hat{\Theta}_k} f^\theta(x_t, x_{t-1}, \dots, x_0)$  denote a maximum likelihood estimator of an agent who is restricted to the class of  $k$ -factor models in  $\hat{\Theta}_k$ . Also define  $h(\theta^*, \theta) = \mathbb{E}^*[\log f^\theta(x_{t+1}|x_t, x_{t-1}, \dots)]$ .

Observe that the KL divergence of any model  $\theta \in \widehat{\Theta}_k$  from the true process can be expressed as follows:  $\text{KL}(\theta^* \|\theta) = h(\theta^*, \theta^*) - h(\theta^*, \theta)$ . As a result,

$$\arg \min_{\theta \in \widehat{\Theta}_k} \text{KL}(\theta^* \|\theta) = \arg \max_{\theta \in \widehat{\Theta}_k} h(\theta^*, \theta),$$

and in particular,

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

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

Next, note that

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

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

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

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

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

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

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

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

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

As a first step, we establish that

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

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 (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,

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

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

Next, note that

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

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

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

and as a result,

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

Taking limits as  $t \rightarrow \infty$  from both sides of the above inequality and using the uniform law of large numbers, it implies that

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

On the other hand, (B.39) guarantees that  $\lim_{t \rightarrow \infty} \text{MSE}(\theta^*, \hat{\theta}_t^{\text{MSE}}) - \text{MSE}(\theta^*, \theta^{\text{KL}}) \geq 0$ . Therefore,

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

Dividing both sides of the above equation by  $\mathbb{E}^*[x_{t+1}^2]$  and using (B.40) and (B.14), then establishes the result.  $\square$

### Proof of Proposition A.3

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

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

for all  $\tau \geq 0$ , where  $\mathbb{E}[\cdot]$  and  $\mathbb{E}^*[\cdot]$  are the subjective expectations of the behavioral and rational agents, respectively, and  $\gamma$  denotes the fraction of rational agents. Consequently,

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

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

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

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

### References

- Allen, Franklin, Stephen Morris, and Hyun Song Shin (2006), “Beauty contests and iterated expectations in asset markets.” *Review of Financial Studies*, 19(3), 719–752.
- Al’pin, Yu. A. (2017), “The Hankel matrix rank theorem revisited.” *Linear Algebra and its Applications*, 534, 97–101.
- Angeletos, George-Marios and Zhen Huo (2021), “Myopia and anchoring.” *American Economic Review*, 111(4), 1166–1200.
- Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry (2020), “Imperfect macroeconomic expectations: Evidence and theory.” In *NBER Macroeconomics Annual* (Martin Eichenbaum and Erik Hurst, eds.), Volume 35, Chapter 1, 1–86, The University of Chicago Press.
- Angeletos, George-Marios and Chen Lian (2018), “Forward guidance without common knowledge.” *American Economic Review*, 108(9), 2477–2512.
- Angeletos, George-Marios and Karthik A. Sastry (2021), “Managing expectations: Instruments vs. targets.” *Quarterly Journal of Economics*, 136, 2467–2532.



- Antoulas, Athanasios C. (2005), *Approximation of Large-Scale Dynamical Systems*. Society for Industrial and Applied Mathematics.
- Aoki, Masanao (1983), "Equivalence of ARMA and state space models." In *Notes on Economic Time Series Analysis: System Theoretic Perspectives*, 22–32, Springer-Verlag, Berlin.
- Bacchetta, Philippe and Eric Van Wincoop (2006), "Can information heterogeneity explain the exchange rate determination puzzle?" *American Economic Review*, 96(3), 552–576.
- Bacchetta, Philippe and Eric van Wincoop (2010), "Infrequent portfolio decisions: A solution to the forward discount puzzle." *American Economic Review*, 100(3), 870–904.
- Bacchetta, Philippe and Eric van Wincoop (2021), "Puzzling exchange rate dynamics and delayed portfolio adjustment." *Journal of International Economics*, 131, 103460.
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer (2015), "X-CAPM: An extrapolative capital asset pricing model." *Journal of Financial Economics*, 115(1), 1–24.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny (1998), "A model of investors sentiment." *Journal of Financial Economics*, 49(3), 307–343.
- Barillas, Francisco and Kristoffer P. Nimark (2017), "Speculation and the term structure of interest rates." *Review of Financial Studies*, 30(11), 4003–4037.
- Berk, Robert H. (1966), "Limiting behavior of posterior distributions when the model is incorrect." *The Annals of Mathematical Statistics*, 37, 51–58.
- Bordalo, Pedro, Nicola Gennaioli, Yueren Ma, and Andrei Shleifer (2020), "Over-reaction in macroeconomic expectations." *American Economic Review*, 110(9), 2748–2782.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018), "Diagnostic expectations and credit cycles." *Journal of Finance*, 73(1), 199–227.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2020), "Memory, attention, and choice." *Quarterly Journal of Economics*, 135(3), 1399–1442.
- Candian, Giacomo and Pierre De Leo (2022), "Imperfect exchange rate expectations." Working paper.
- Chinn, Menzie D. and Jeffrey A. Frankel (2020), "A third of a century of currency expectations data: The carry trade and the risk premium." Working Paper, Harvard University.
- Coibion, Olivier and Yuriy Gorodnichenko (2015), "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review*, 105(8), 2644–2678.
- da Silveira, Rava Azeredo, Yeji Sung, and Michael Woodford (2020), "Optimally imprecise memory and biased forecasts." NBER Working Paper No. 28075.

- Donald, Stephen G., Natércia Fortuna, and Vladas Pipiras (2007), “On rank estimation in symmetric matrices: The case of indefinite matrix estimators.” *Econometric Theory*, 23(6), 1217–1232.
- Engel, Charles (2016), “Exchange rates, interest rates, and the risk premium.” *American Economic Review*, 106(2), 436–474.
- Esponda, Ignacio and Demian Pouzo (2016), “Berk–Nash equilibrium: A framework for modeling agents with misspecified models.” *Econometrica*, 84(3), 1093–1130.
- Fama, Eugene F. (1984), “Forward and spot exchange rates.” *Journal of Monetary Economics*, 14(3), 319–338.
- Farhi, Emmanuel and Iván Werning (2019), “Monetary policy, bounded rationality, and incomplete markets.” *American Economic Review*, 109(11), 3887–3928.
- Farmer, Leland E., Emi Nakamura, and Jón Steinsson (2022), “Learning about the long run.” NBER Working Paper 29495.
- Froot, Kenneth A. and Jeffrey A. Frankel (1989), “Forward discount bias: Is it an exchange risk premium?” *Quarterly Journal of Economics*, 104(1), 139–161.
- Fuster, Andreas, Benjamin Hébert, and David Laibson (2012), “Natural expectations, macroeconomic dynamics, and asset pricing.” In *NBER Macroeconomics Annual* (Daron Acemoglu and Michael Woodford, eds.), Volume 26, Chapter 1, 1–48, The University of Chicago Press.
- Fuster, Andreas, David Laibson, and Brock Mendel (2010), “Natural expectations and macroeconomic fluctuations.” *Journal of Economic Perspectives*, 24(4), 67–84.
- Gabaix, Xavier (2014), “A sparsity-based model of bounded rationality.” *Quarterly Journal of Economics*, 129(4), 1661–1710.
- Gourinchas, Pierre-Olivier and Aaron Tornell (2004), “Exchange rate puzzles and distorted beliefs.” *Journal of International Economics*, 64(2), 303–333.
- Granziera, Eleonora and Markus Sihvonen (2022), “Bonds, currencies and expectational errors.” Norges Bank Research Paper 2020/3.
- Hespanha, João P. (2018), *Linear Systems Theory*. Princeton University Press.
- Hirshleifer, David, Jun Li, and Jianfeng Yu (2015), “Asset pricing in production economies with extrapolative expectations.” *Journal of Monetary Economics*, 76, 87–106.
- Horn, Roger A. and Charles R. Johnson (2013), *Matrix Analysis*, second edition. Cambridge University Press, New York, NY.
- Kalemli-Özcan, Şebnem and Liliana Varela (2021), “Five facts about the uip premium.” NBER Working Paper.

- Kohlhas, Alexandre N. and Ansgar Walther (2021), “Asymmetric attention.” *American Economic Review*, 111(9), 2879–2925.
- Molavi, Pooya (2022), “Simple models and biased forecasts.” Working Paper, Northwestern University.
- Nagel, Stefan and Zhengyang Xu (2022), “Asset pricing with fading memory.” *Review of Financial Studies*, 35(5), 2190–2245.
- Newey, Whitney K. and Kenneth D. West (1987), “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.” *Econometrica*, 55(3), 703–708.
- Panageas, Stavros (2020), “The implications of heterogeneity and inequality for asset pricing.” *Foundations and Trends in Finance*, 12(3), 199–275.
- Piazzesi, Monika, Juliana Salomao, and Martin Schneider (2015), “Trend and cycle in bond premia.” Working paper.
- Rabin, Matthew and Dimitri Vayanos (2010), “The gambler’s and hot-hand fallacies: Theory and applications.” *Review of Economic Studies*, 77(2), 730–778.
- Sandberg, Henrik (2019), “Introduction to model order reduction.” Lecture notes, <http://www.diva-portal.org/smash/get/diva2:1338145/fulltext01.pdf>.
- Shalizi, Cosma Rohilla (2009), “Dynamics of Bayesian updating with dependent data and misspecified models.” *Electronic Journal of Statistics*, 3, 1039–1074.
- Sims, Christopher A. (2003), “Implications of rational inattention.” *Journal of Monetary Economics*, 50(3), 665–690.
- Stavrakeva, Vania and Jenny Tang (2020), “A fundamental connection: Survey-based exchange rate decomposition.” Working paper.
- Valchev, Rosen (2020), “Bond convenience yields and exchange rate dynamics.” *American Economic Journal: Macroeconomics*, 12(2), 124–166.
- Wachter, Jessica A. and Michael Jacob Kahana (2021), “A retrieved-context theory of financial decisions.” NBER Working Paper No. 26200.

# Online Appendix

## C Technical Appendix

### C.1 Lemmas for the Proof of Theorem 1

In this appendix, we state and prove two lemmas that are used in the proof of Theorem 1.

**Lemma C.1.** *Let  $\mathbf{M}$  and  $u$  be defined as in (B.11) and let  $\hat{\sigma}_x^2 = c' \hat{\Sigma} c$ . Minimizing (B.13) with respect to  $(\mathbf{M}, u, \hat{\sigma}_x^2)$  and subject to the constraint that  $\rho(\mathbf{M}(\mathbf{I} - uu')\mathbf{M}') \leq 1$  is equivalent to minimizing (B.10) with respect to  $(\mathbf{A}, \mathbf{B}, c)$ .*

*Proof.* In the proof of Theorem 1, we showed that for any  $(\mathbf{A}, \mathbf{B}, c)$ , there exists  $(\mathbf{M}, u, \hat{\sigma}_x^2)$  such that the values of (B.10) and (B.13) coincide. Furthermore, (B.11) implies that  $\mathbf{M}$  and  $u$  satisfy  $\rho(\mathbf{M}(\mathbf{I} - uu')\mathbf{M}') \leq 1$ . Therefore, it is sufficient to show that for any  $(\mathbf{M}, u, \hat{\sigma}_x^2)$  such that  $\rho(\mathbf{M}(\mathbf{I} - uu')\mathbf{M}') \leq 1$ , there always exists  $(\mathbf{A}, \mathbf{B}, c)$  such that the values of (B.10) and (B.13) coincide. To this end, let  $(\mathbf{M}, u, \hat{\sigma}_x^2)$  be an arbitrary tuple such that  $\rho(\mathbf{M}(\mathbf{I} - uu')\mathbf{M}') \leq 1$  and define  $(\mathbf{A}, \mathbf{B}, c)$  as follows:

$$\mathbf{A} = \mathbf{M}, \quad \mathbf{B} = \hat{\sigma}_x(\mathbf{I} - \mathbf{M}(\mathbf{I} - uu')\mathbf{M}')^{1/2}, \quad c = u.$$

Given these definitions, it is immediate to verify that the solution to the algebraic Riccati equation (B.7) and the corresponding Kalman gain in (B.8) are given by  $\hat{\Sigma} = \hat{\sigma}_x^2 \mathbf{I}$  and  $g = \mathbf{M}u$ , respectively. This also implies that  $\mathbf{M}(\mathbf{I} - uu') = \mathbf{A} - gc'$ . Replacing for  $\mathbf{M}(\mathbf{I} - uu')$ ,  $\mathbf{M}u$ , and  $u'$  on the right-hand side of (B.13) by, respectively,  $\mathbf{A} - gc'$ ,  $g$ , and  $c'$  leads to the right-hand side of (B.10).  $\square$

The next lemma is used to prove the claim in footnote 26 in the proof of Theorem 1.

**Lemma C.2.** *Suppose  $v, q \in \mathbb{R}^k$  satisfy  $v'q > 0$ . Then, there exists a symmetric, positive definite matrix  $\mathbf{R}$  such that  $\mathbf{R}v = q$ .*

The claim in footnote 26 follows immediately from the above lemma. Specifically, since  $c' \hat{\Sigma} c > 0$ , setting  $v = c$  and  $q = \hat{\Sigma} c$  in the above lemma implies that there exists a symmetric, positive definite matrix  $\mathbf{R}$  such that  $\mathbf{R}c = \hat{\Sigma} c$ .

*Proof.* If  $v$  and  $q$  are parallel, then setting  $\mathbf{R}$  as a positive multiple of the identity matrix,  $\mathbf{R} = (\|q\|_2 / \|v\|_2) \mathbf{I}$ , immediately implies that  $q = \mathbf{R}v$ . Therefore, in the rest of the proof, we assume that  $v$  and  $q$  are not parallel.

Let  $\text{span}(q, v)$  denote the 2-dimensional subspace of  $\mathbb{R}^k$  spanned by  $q$  and  $v$ . Since  $q$  and  $v$  are neither parallel nor orthogonal, there exist unit-norm vectors  $u_1, u_2 \in \text{span}(q, v)$  such that  $u_1' u_2 = 0$  and  $u_1' v, u_1' q, u_2' v, u_2' q > 0$ . Given  $u_1$  and  $u_2$ , let  $\{u_3, \dots, u_k\}$  be a set of vectors such that  $\{u_i\}_{i=1}^k$  forms an orthonormal basis in  $\mathbb{R}^k$ . Note that, by construction,  $u_i' v = 0$  for  $i = 3, 4, \dots, k$ . Define

$$\mathbf{R} = \sum_{i=1}^n \alpha_i u_i u_i',$$

where  $\{\alpha_1, \dots, \alpha_k\}$  is a collection of strictly positive numbers. Clearly, matrix  $\mathbf{R}$  is positive definite. Therefore, in what follows, we show that one can choose  $\{\alpha_1, \dots, \alpha_k\}$  such that  $\mathbf{R}v = q$ . On the one hand, note that,

$$\mathbf{R}v = \alpha_1(u'_1v)u_1 + \alpha_2(u'_2v)u_2.$$

On the other hand, since  $q \in \text{span}(q, v) = \text{span}(u_1, u_2)$ ,

$$q = (u'_1q)u_1 + (u'_2q)u_2.$$

Comparing the above equations therefore implies that  $\mathbf{R}v = q$  as long as  $\alpha_1 = (u'_1q)/(u'_1v)$  and  $\alpha_2 = (u'_2q)/(u'_2v)$ . Note that  $\alpha_1$  and  $\alpha_2$  are strictly positive. Setting  $\alpha_3, \dots, \alpha_k$  to be arbitrary strictly positive scalars, then implies that  $\mathbf{R}v = q$ , where  $\mathbf{R}$  is positive definite.  $\square$

## C.2 Over- and Under-Reaction to Information

Recall from Proposition 3 that when the set of models entertained by agents is not rich enough to contain the true data-generating process (i.e., when  $k < n$ ), they end up with subjective expectations that generate return predictability. This departure from rational expectations is a consequence of the fact that, due to their misspecified model of the data-generating process, agents do not incorporate new pieces of information into their forecasts as they would have had they known the true process.

We show that whether agents over- or under-react to new information cannot be decoupled from the environment they live in. In particular, we argue that agents' forecasts may exhibit systematic over- or under-reaction to news depending on (i) the statistical properties of the underlying data-generating process and (ii) the horizon of interest. This means that, in our framework, over- and under-reaction of expectations are endogenous and are not baked into agents' expectations formation process.

To measure the extent of over- and under-reaction to new information at different horizons, we follow Coibion and Gorodnichenko (2015) and consider the family of regressions

$$x_{t+h} - \mathbb{E}_t[x_{t+h}] = \alpha_h^{\text{CG}} + \beta_h^{\text{CG}}(\mathbb{E}_t[x_{t+h}] - \mathbb{E}_{t-1}[x_{t+h}]), \quad (\text{C.1})$$

where, as before,  $x_t$  is the realization of the fundamental at time  $t$  and  $\mathbb{E}[\cdot]$  denotes agents' subjective expectations. The regressand on left-hand side of (C.1) is agents' forecast error for realization of the fundamental  $h$  periods in the future, while the regressor on the right-hand side is their latest forecast revision. Thus,  $\beta_h^{\text{CG}} > 0$  means that when agents revise their  $h$ -step-ahead forecast of the fundamental upward at time  $t$ , they tend to undershoot its eventual realization at time  $t+h$ . In other words, agents systematically under-react to new information at time  $t$ . By a similar logic,  $\beta_h^{\text{CG}} < 0$  means that agents tend to systematically over-react to new information when forming expectations about the realization of the fundamental  $h$  periods in the future. As is well known, under rational expectations (i.e., when  $\mathbb{E}[\cdot] = \mathbb{E}^*[\cdot]$ ), forecast revisions are orthogonal to future forecast errors, which means that  $\beta_h^{\text{CG}} = 0$  for all horizons  $h \geq 1$ .

Turning to our framework, suppose that agents are constrained to using single-factor models (i.e.,  $k = 1$ ) to make forecasts about the future realization of the fundamental, even though the

true data-generating process may be driven by  $n \neq k$  factors. Under such a specification, the slope coefficient of regression (C.1) at horizon  $h$  is given by

$$\beta_h^{\text{CG}} = \frac{\xi_h^* - \xi_1^* \xi_{h+1}^*}{\xi_1^{*h} (1 - \xi_1^{*2})} - 1, \quad (\text{C.2})$$

where  $\xi_h^*$  denotes the autocorrelation of the fundamental at lag  $h$ .

Two observations are immediate. First, whether agents over- or under-react to new information depends on the statistical properties of the underlying data-generating process. For example, it is straightforward to verify that  $\beta_1^{\text{CG}} > 0$  if  $\xi_1^{*2} > \xi_2^*$ , whereas  $\beta_1^{\text{CG}} < 0$  if  $\xi_1^{*2} < \xi_2^*$ . Thus, the same agent with the same level of sophistication may end up with forecasts that under- or overshoot the eventual realization of the fundamental depending on the environment. Second, holding the underlying data-generating process fixed, the sign and the magnitude of the right-hand side of (C.2) varies with  $h$ . This means that agents' expectations may over-react to new information at some horizon, while simultaneously under-reacting at others. As already discussed, the exact pattern of over- and under-reaction depends on the process that governs the fundamental.

**Derivation of Equation (C.2).** Since  $k = 1$ , agents' model can be represented by the tuple of scalars  $\theta = (a, b, c)$ , where  $a \in (-1, 1)$  denotes the persistence parameter of the underlying factor. This means that  $\mathbb{E}_t[x_{t+h}] = a^h x_t$  for all  $h \geq 0$ . As a result, the slope coefficient of the regression in (C.1) is given by

$$\beta_h^{\text{CG}} = \frac{\mathbb{E}^*[(x_{t+h} - a^h x_t)(x_t - a x_{t-1})]}{a^h \mathbb{E}^*[(x_t - a x_{t-1})^2]} = \frac{\xi_h^* - a \xi_{h+1}^* - a^h (1 - a \xi_1^*)}{a^h (1 + a^2 - 2a \xi_1^*)}, \quad (\text{C.3})$$

where  $\xi_h^*$  denotes the autocorrelation of the fundamental at lag  $h$ . Next, recall from the proof of Proposition 6 that when agents are restricted to single-factor models, the objective function in (12) reduces to (B.23), in which  $|m| < 1$  is a scalar. Optimizing (B.23) over  $m$  implies that  $m = \xi_1^*$ . Therefore, by equation (B.11), the persistence parameter is given by  $a = \xi_1^*$ . Plugging this expression into (C.3) establishes (C.2).

### C.3 Micro-Founded Model for Heterogenous-Agent Economy

As discussed in Appendix A.3, Proposition A.3 applies to any asset pricing framework that has a reduced-form representation in the form of equation (A.3). In this appendix, we present a stylized micro-founded model with such a property. The model, which is a variant of the heterogenous-agent economy of Allen, Morris, and Shin (2006), has a reduced-form representation identical to equation (A.3) and reduces to equation (2) when all agents are identical.<sup>28</sup>

Consider a discrete-time economy with a single consumption good and a single risky asset. The risky asset is in zero net supply and delivers a dividend stream  $\{x_t\}_{t=-\infty}^{\infty}$  of the consumption good. The economy is populated by overlapping generations of agents, who live for 2 periods. A new generation of agents of unit mass is born at each date  $t$ . Agents in the same generation have

<sup>28</sup>Also see Bacchetta and Van Wincoop (2006) for a similar model.

identical preferences and can observe the entire past realizations of  $x_t$  but may have heterogeneous subjective expectations about the process that drives the asset’s dividend stream (and hence, have different expectations about future dividends). Specifically, we assume that a fraction  $1 - \gamma$  of agents are restricted to using models consisting of at most  $k$  factors, while the remaining  $\gamma$  fraction are unconstrained and can entertain models with any number of factors.

All agents only consume when they are old. When young, agents build up a position in the asset, which they then unwind in the next period—when they are old—to acquire the consumption good. The utility of an agent  $i \in [0, 1]$  who is born in period  $t$  and acquires  $q_{it}$  units of the asset is given by

$$u_i(q_{it}) = q_{it}(x_t - y_t + y_{t+1}) - \frac{1}{2}\gamma q_{it}^2,$$

where  $y_t$  denotes the price of the asset at time  $t$  in units of the consumption good and  $\gamma > 0$  parameterizes a quadratic cost of trading the asset.<sup>29</sup>

Given the above, it is immediate that the demand for the asset by an agent  $i$  born at time  $t$  is given by  $q_{it} = \frac{1}{\gamma}(x_t - y_t + \mathbb{E}_{it}[y_{t+1}])$ , where  $\mathbb{E}_{it}[\cdot]$  denotes the agent’s subjective expectations. Consequently, market clearing implies that the price of the asset at time  $t$  satisfies

$$y_t = x_t + \int_0^1 \mathbb{E}_{it}[y_{t+1}] di.$$

In other words, the price of the asset at time  $t$  is equal to its dividend plus the cross-sectional average of agents’ subjective expectations of the price the next period. Given our assumption that only a fraction  $1 - \gamma$  of the agents are subject to our behavioral constraint, this implies that  $y_t = x_t + \bar{\mathbb{E}}_t[y_{t+1}]$ , where  $\bar{\mathbb{E}}[\cdot] = \gamma\mathbb{E}^*[\cdot] + (1 - \gamma)\mathbb{E}[\cdot]$ .

## D Data Appendix

In this appendix, we provide a brief description of how we construct the realized forecast errors of interest rate differentials used in the analysis in Subsection 5.3.

We obtain interest rate forecasts from *Consensus Economics* between June 1998 and December 2021 for the following set of countries: Canada, the Eurozone, Japan, Switzerland, the United Kingdom, and the United States. 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. Using these surveys, we construct a “consensus forecast” by calculating the median forecast for each country. This results in a monthly sample for 3-month-ahead consensus forecasts of 3-month interest rates in each of the six countries. We restrict our sample to the June 1998–December 2021 period in order to have a non-trivial cross-section of different forecasts in each country.

To measure the realized counterparts to the survey forecasts, we obtain the following 3-month interest rates:

---

<sup>29</sup>As in Allen, Morris, and Shin (2006), the assumption of overlapping generations of agents is intended to abstract from dynamic incentives. The quadratic trading costs are imposed to ensure that disagreement among agents does not lead to unbounded trade in equilibrium.

- Canada: 3-Month Treasury Bill rate available from Bank of Canada webpage;
- Eurozone: 3-Month or 90-day Rates and Yields: Interbank Rates available from FRED;
- Japan: 3-Month Yen Certificates of Deposit until June 2010; 3-month Yen TIBOR Rate from July 2010, available from the Bank of Japan webpage;
- Switzerland: 3-Month or 90-day Rates and Yields: Eurodollar Deposits available from FRED;
- United Kingdom: 3-Month or 90-day Rates and Yields: Interbank Rates available from FRED;
- United States: 3-Month Treasury Bill Secondary Market Rate available from FRED.

We then construct the forecast errors for each interest rate by subtracting the survey forecasts from their realized counterparts. We follow a similar procedure to construct forecast errors for interest rate differentials.

## E Cross-Sectional Evidence

In the application on UIP violations in Section 5, we followed Engel (2016) and focused on trade-weighted average exchange rates and interest rate differentials of various currencies vis-à-vis the United States. In this appendix, we redo the exercises in Subsection 5.2 and 5.3 for each of the currencies in our sample separately.

### E.1 Return Predictability

We calculate the model-implied slope coefficients of the return predictability regression (6) for each currency separately (all against the U.S. dollar) and plot them against the corresponding coefficients estimated from the data. Figure E.1 depicts the results for  $k = 1$  at different horizons. Recall from Theorem 1 and Proposition 6 that model-implied slope coefficients depend on the shape of the autocorrelation function of the fundamental (in this case, the interest rate differential between the corresponding country and the U.S.). Therefore, as the autocorrelations differ in the cross-section of countries, so should the model-implied slope coefficients,  $\beta_h^{\text{FX}}$ . Nonetheless, as Figure E.1 illustrates, the patterns of model-implied slope coefficients closely track those of the coefficients estimated from return data for each currency. This relationship holds both at short horizons when most coefficients are positive, as well as for the longer horizons. Once again, we emphasize that we do not use exchange rate and return data for generating any of the model-implied slope coefficients. Finally, note that, as in Figure 2 for the trade-weighted basket of currencies, the model-implied coefficients in Figure E.1 underestimate the magnitudes of their empirical counterparts.

### E.2 Forecast-Error Predictability

We also reproduce the results on forecast-error predictability in Subsection 5.3 separately for each of the countries in our sample. To estimate the coefficients of forecast-error predictability regressions,



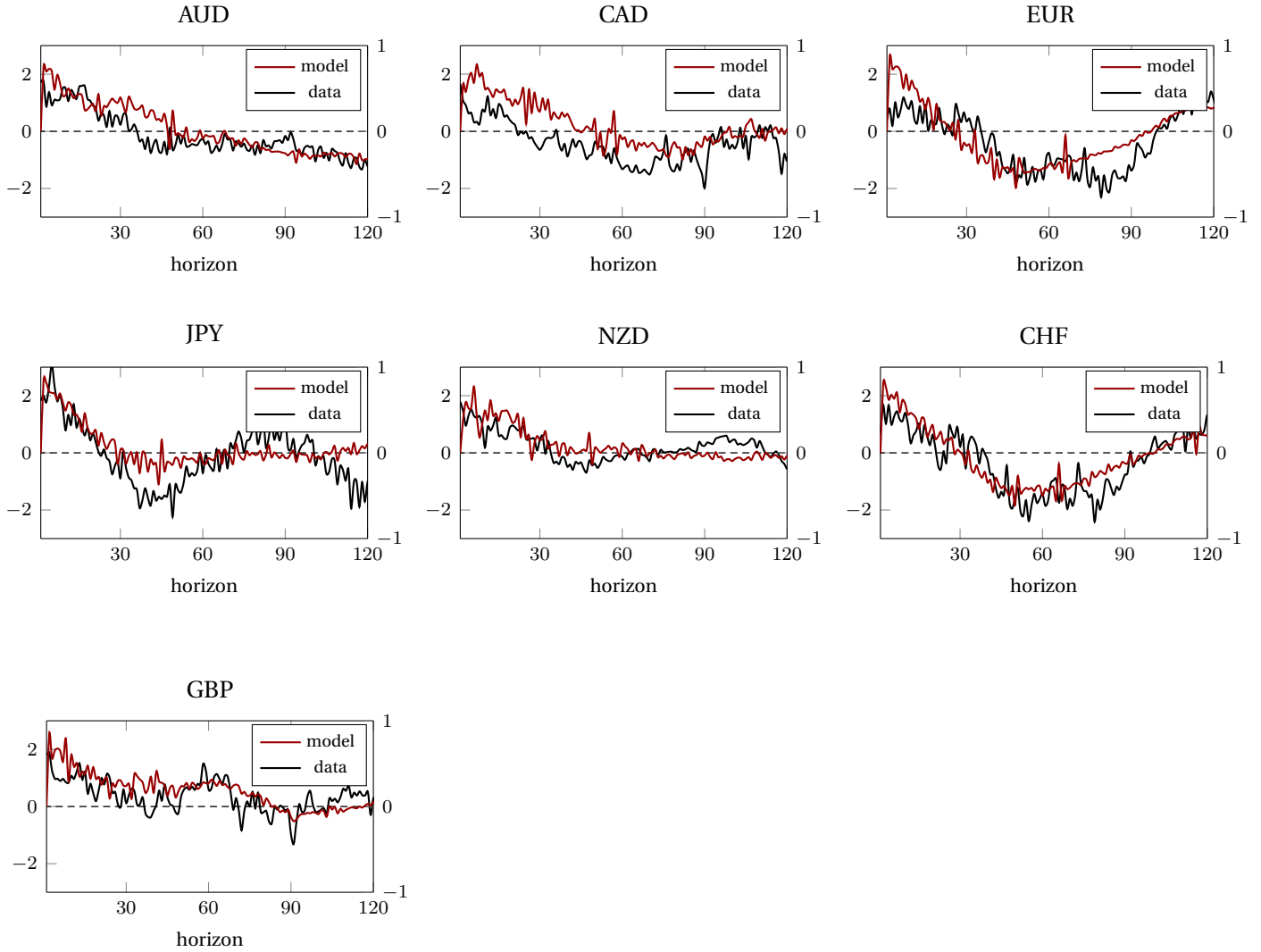


Figure E.1. Return Predictability in the Cross-Section

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

we regress the forecast error of each interest rate differential relative to the U.S. (constructed from *Consensus Economics* data) on the corresponding realized interest rate differential. We also obtain the model-implied slope coefficients of the forecast-error predictability regression for each country from the corresponding autocorrelation of the interest rate differential process under the assumption that  $k = 1$  using equation (16). Once again, note that, as the autocorrelations differ in the cross-section of countries, so should the model-implied slope coefficients,  $\beta_h^{\text{fe}}$ . Figure E.2 depicts the results.

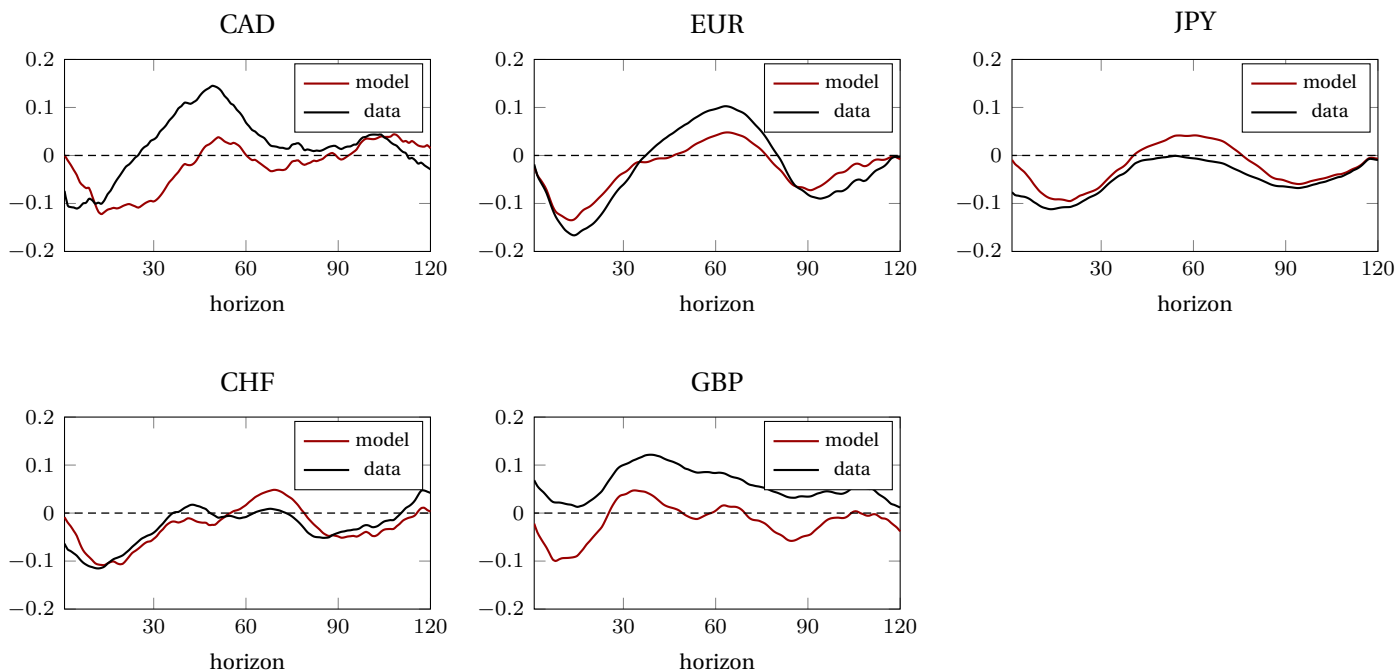


Figure E.2. Forecast-Error Predictability in the Cross-Section

*Notes:* This figure plots estimated slope coefficients of forecast-error predictability regression of the interest rate differential process (vis-à-vis the U.S.) for each country together with the corresponding model-implied coefficients from a one-factor model. Monthly data from June 1998 to December 2021.

### E.3 Testing for $k$

Finally, we use our result in Proposition 8 to directly test for the value of  $k$  for each of the countries in our sample separately. We follow the same 3-step procedure laid out in Subsection 5.3. For each country, we (i) regress the consensus (3-month-ahead) forecasts of interest rate differential vis-à-vis the United States obtained from survey data on the past realizations of the interest rate differential to estimate coefficients  $(\phi_1^{(3)}, \phi_2^{(3)}, \dots)$  in (17); (ii) construct the Hankel matrix  $\Phi$  with typical element  $\Phi_{ij} = \phi_{i+j-1}^{(3)}$ ; and (iii) test for the rank of  $\Phi$ . 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. Additionally, note that step (iii) of the estimation procedure requires testing for the rank of an infinite-dimensional matrix. This, however, is not an issue for our empirical implementation, as by Kronecker's Theorem (Al'pin, 2017, Theorem 2),  $\text{rank}(\Phi)$  is equal to the rank of the finite-dimensional Hankel matrix constructed from coefficients  $(\phi_1^{(3)}, \dots, \phi_s^{(3)})$  for any odd integer  $s$  such that  $s \geq 2k - 1$ . In what follows, we set  $s = 7$ , but find the same result for other values of  $s$ .

Table E1 reports the  $p$ -values corresponding to the rank test of Donald, Fortuna, and Pipiras (2007) for the null hypothesis that  $\text{rank}(\Phi) = 1$  for different countries (including the trade-weighted average studied in Subsection 5.3) and different lags in equation (17).

Table E1. Testing for  $k$

No. of lags	Country					Average
	CAD	EUR	JPY	CHF	GBP	
10	0.603	0.938	0.875	0.923	0.182	0.151
11	0.590	0.954	0.185	0.936	0.176	0.986
12	0.547	0.970	0.325	0.693	0.160	0.883
13	0.419	0.974	0.616	0.003	0.178	0.081
14	0.295	0.980	0.607	0.943	0.098	0.974
15	0.266	0.964	0.582	0.906	0.072	0.953
16	0.097	0.978	0.648	0.973	0.069	0.859
17	0.084	0.983	0.727	0.981	0.068	0.977
18	0.080	0.968	0.714	0.881	0.080	0.963
19	0.086	0.971	0.664	0.844	0.097	0.949
20	0.141	0.979	0.648	0.794	0.091	0.916

*Notes:* This table reports  $p$ -values of the rank test of [Donald, Fortuna, and Pipiras \(2007\)](#) for the null hypothesis that  $\text{rank}(\Phi) = 1$  for different countries. Matrix  $\Phi$  is a Hankel matrix constructed from coefficients in (17). The first column contains the number of lags on the right-hand side of (17).