Estimation and Comparison of Beta-Pricing Models *

Siddhartha Chib Guanhao Feng Jingyu He Qianshu Zhang

First draft: Jan. 2025; this draft: Feb. 2025

Abstract

We develop a Bayesian framework for estimating beta-pricing models, with traded and nontraded factors, that integrates key theory-implied asset pricing constraints and, crucially, facilitates model comparisons through marginal likelihoods. Simulations show that our marginal likelihood based model comparisons outperform traditional metrics such as cross-sectional R^2 and Hansen-Jagannathan distance. Empirically, we find that most nontraded factors are weak factors, except for the intermediary capital ratio (He, Kelly, and Manela, 2017). Given a set of factors, the economic performance of the top model is validated by the out-of-sample tangency portfolio Sharpe Ratios. Upon revisiting the findings of Kan, Robotti, and Shanken (2013), we find that adding nondurable consumption growth enhances the Fama-French 3-factor model. Analyzing consumption-based models, we find that various consumption growth measures are weak, although garbage consumption growth appears to be the most useful.

Key Words: Asset Pricing Constraint, Marginal Likelihood, Model Comparison, Nontraded Factors, Risk Premia.

*We thank Raymond Kan, Semyon Malamud, Gustavo Schwenkler, Chu Zhang, Lingxiao Zhao, Guofu Zhou, and the seminar and conference participants at the City University of Hong Kong for their valuable comments. Chib (E-mail: chib@wustl.edu) is at Washington University in St. Louis; Feng (E-mail: gavin.feng@cityu.edu.hk), He (E-mail: jingyuhe@cityu.edu.hk), and Zhang (E-mail: qszhang7-c@my.cityu.edu.hk) are at City University of Hong Kong.

We do not have disclosure of any sources of research support and any financial relationships or other potential conflicts of interest.

1 Introduction

Beta-pricing models, including the Capital Asset Pricing Model (CAPM), provide a foundational framework for explaining cross-sectional average returns by linking asset returns to their loadings (betas) on systematic risk factors. These models are integral to portfolio construction, aligning investments with factor exposures, assessing asset return sensitivities to systematic risks, and decomposing returns into systematic and idiosyncratic components. Beta-pricing frameworks, such as Breeden et al. (1989) and Fama and French (1993), are cornerstones of financial theory and practice, continually evolving as researchers identify new risk factors and enhance methods for estimating and testing asset pricing models.

Unlike other linear factor models, beta-pricing models integrate both asset-specific time-series dynamics and cross-sectional asset pricing restrictions, offering a more comprehensive framework for understanding asset returns. Estimation methods include the two-pass regression approach, maximum likelihood estimation (MLE), and the generalized method of moments (GMM). Despite their widespread use, estimating factor risk premia and identifying the relevant set of risk factors remains challenging due to issues such as weak factors, model misspecification, and errors-in-variables (e.g., Kleibergen, 2009; Gospodinov et al., 2019; Giglio et al., 2024).

Since no asset pricing model has achieved universal acceptance or successfully withstood rejection under rigorous asset pricing tests (e.g., Fama and French, 2015; Hou et al., 2021), evaluating the relative pricing performance of competing models remains a critical endeavor. For models that rely solely on traded factors, performance comparisons can be conducted without dependence on test assets (e.g., Fama and French, 2018; Barillas and Shanken, 2018; Chib, Zeng, and Zhao, 2020).

However, models incorporating both traded and nontraded factors face additional challenges. These models are typically evaluated using two key metrics: cross-sectional R^2 (CSR^2) and Hansen-Jagannathan (H-J) distance (Hansen and Jagannathan, 1997), which summarize quadratic pricing errors. The CSR^2 captures how well the

model explains variation in average returns, while the H-J distance measures pricing errors from the moment conditions. However, both metrics have limitations, as larger models often outperform smaller ones. Although asymptotic distributions for these metrics have been derived (see, Kan et al., 2013; Kan and Robotti, 2009), enabling statistical tests to compare competing models, these tests frequently exhibit low power. Consequently, even when the null hypothesis of equality cannot be rejected, the choice of the preferred model remains ambiguous. Recent advancements, such as Feng et al. (2020) and Bryzgalova et al. (2023), have improved model selection techniques in cross-sectional regression but do not focus on risk premia. This underscores the persistent need for more robust and reliable model evaluation methods.

From the Stochastic Discount Factor (SDF) framework, we know that the risk premium of a traded factor equals its mean if and only if it is a priced risk factor (e.g., Cochrane, 2009). However, this fundamental economic constraint is frequently not imposed in empirical studies employing two-pass cross-sectional regression with known risk factors. As a result, some studies report substantial zero-beta rate estimates that frequently exceed the associated factor risk premia or, in extreme instances, result in negative market risk premia (e.g., Kan et al., 2013). These distortions can diminish the statistical power of tests designed to evaluate risk factors and asset pricing models, thereby compromising the reliability of empirical findings.

To address these challenges, we develop a framework to identify which subset of the available factors are priced risk factors. Our approach builds on the standard beta-pricing model, which incorporates the $\boldsymbol{B\lambda}$ cross-sectional pricing condition (e.g., Cochrane, 2009)[1]. We extend this formulation in three key ways. First, we distinguish between traded $\boldsymbol{f^T}$ and nontraded $\boldsymbol{f^{NT}}$ risk factors and refine the cross-sectional pricing condition as $\boldsymbol{B\lambda} = \boldsymbol{B^T\lambda^T} + \boldsymbol{B^{NT}\lambda^{NT}}$. Second, we impose the restriction that for traded risk factors, $\boldsymbol{\lambda^T} = E[\boldsymbol{f^T}]$. We refer to the resulting model for excess returns as

[1]We adopt a Bayesian approach, avoiding the standard two-pass regression, which uses first-pass beta estimates as inputs in the second pass. Instead, we incorporate the second pass as a cross-sectional constraint within the overall likelihood. The standard two-pass method assumes separate likelihoods for the same data, and treating first-pass estimates as data in the second pass misrepresents parameter uncertainty from a Bayesian perspective. See Section 2.6 for further discussion.

the *economically restricted time series model*, or simply the *return model*. Third, to enable model comparisons of return models with different sets of traded and nontraded risk factors, we augment the latter model for excess returns with a model that accounts for the entire pool of available factors, which we call the *factor model*.

We estimate our models using Bayesian methods, with black-box, training sample priors that require no user intervention. Our approach provides a principled framework for comparing beta-pricing models with different risk factors in the return model. Models are compared using marginal likelihoods, which we estimate efficiently by the method of Chib (1995). Since, in large samples, the model with the highest marginal likelihood is either the true model, or the closest approximation to it in terms of Kullback-Leibler distance (Chib et al., 2018), we infer the set of best risk factors from the maximal marginal likelihood model.

A key motivation for our approach is the challenge posed by weakly correlated factors, which can lead to unreliable risk premium estimates when using traditional methods such as two-pass regression, MLE, and GMM. While prior studies have focused on test statistics to detect factor weakness (e.g., Kleibergen and Zhan, 2020, F test of rank condition), our model comparison approach offers a different route. Rather than testing individual beta coefficients, we assess overall model fit by comparing models that include weak factors to those that exclude them, with the comparison being made through the marginal likelihood that integrates out the beta parameters. Through a simulation study, we demonstrate that incorporating a weak factor reduces the marginal likelihood relative to a benchmark model that omits it, even when the estimated risk premium appears significant.

We conduct several simulation studies to validate the effectiveness and robustness of our proposed methodology. First, we utilize a simulation framework akin to that of Lewellen et al. (2010) to evaluate the significance of imposing the traded factor constraint in a model comprising only traded factors. Next, we conduct two simulations to assess both the estimation of risk premia in a correctly specified model and the performance of model selection across approximately 1,000 specifications. Un-

der the correct model specification, our Bayesian risk premia estimates exhibit lower bias and root mean square error (RMSE) compared to those obtained using the two-pass weighted least squares (WLS) approach. Furthermore, the model that maximizes marginal likelihood effectively identifies the true model. In contrast, we also show that CSR^2 and H-J distance fail to distinguish the optimal model from misleading alternatives, highlighting the limitations of conventional model evaluation metrics.

Empirical Highlights. We analyze monthly U.S. equity data from 1985 to 2023, examining eight traded and eight nontraded factors using a broad cross-section of 302 equity portfolios. The traded factors include the six Fama-French factors (Fama and French, 2018), augmented by BAB and QMJ. The nontraded factors span industrial production growth, equity market liquidity, long-term interest rates, M2, durable and nondurable consumption growth, intermediary capital, and presidential approval ratings. We also revisit nontraded factor models from Kan et al. (2013) based on CSR^2, alongside the consumption-based models tested in Kleibergen and Zhan (2020), employing identical data and sample periods.

First, we test each nontraded factor and evaluate their incremental contributions relative to benchmark traded factor models using marginal likelihood. We find that most of these factors appear to be weak. For instance, LTY marginally outperforms the CAPM but underperforms compared to other models. Similarly, PCEDG and PEAR outperform the Fama-French three-factor model (FF3) but do not perform as well against stronger benchmarks. Among all these nontraded factors, only HKMcapital consistently outperforms the benchmark models, suggesting its value as a nontraded factor, with a risk premium estimate of 1.4%, which is close to the average return of its traded counterpart.

Second, among the 32,768 model specifications derived from the 16 factors, the best-performing model includes all traded factors and HKMcapital, forming a concise nine-factor framework. The second-best model consists of all traded factors without any nontraded factor, emphasizing the overall weakness of most nontraded factors. In the optimal model, the risk premium estimate for HKMcapital is significant across

Bayesian methods, WLS, and mimicking portfolio approaches, underscoring the statistical robustness of the model.

In addition, we assess model performance using the tangency portfolio Sharpe Ratios of test assets. By selecting the best model in-sample and utilizing the estimated model-implied mean and covariance for tangency portfolio weight, we find that the tangency portfolio of test assets constructed with the top model outperforms not only in-sample but also out-of-sample. Bayesian estimation also outperforms WLS when dealing with a large number of test assets. This highlights the economic significance of model comparison and estimation using our Bayesian approach.

Although Kan et al. (2013) find that the ICAPM model is the best among the eight models they compare, we find that the FF3 model yields the highest marginal likelihood. When we further consider all combinations of 14 factors, the best-performing model integrates the FF3 with the nondurable consumption growth factor, with FF3 ranking as the second-best model overall. Revisiting the consumption-based model presented in Kleibergen and Zhan (2020), we find that when the market factor is included as a test asset in our empirical analysis, the garbage consumption growth measure outperforms all other measures. However, all consumption measures are deemed weak when controlling for MKTRF as a traded factor.

Related Literature. This paper contributes to the literature on asset pricing model comparisons. In the context of traded factor models, Gibbons, Ross, and Shanken (1989, henceforth the GRS test) link the Sharpe ratio of the factor-efficient portfolio to cross-sectional pricing errors. Recently, Fama and French (2018) propose the maximum squared Sharpe ratio metric, Barillas et al. (2020) examine differences in squared Sharpe ratios, and Kan et al. (2024) extend this framework to incorporate estimation risk in Sharpe ratio comparisons.

Beta-pricing models with traded and non trade factors have been analyzed using various approaches, including Kan and Robotti (2009), which examines the H-J distance as an alternative measure of model fit, and Gospodinov et al. (2014), which extends this framework to address model misspecification. Additionally, Kan et al.

(2013) derive the asymptotic distribution of CSR^2 and use it to evaluate model performance. Although these methods provide effective benchmarks against alternatives, they do not identify the optimal model specification.

Furthermore, Lewellen et al. (2010) critique the reliance on CSR^2 and offer several recommendations to enhance empirical tests of pricing models. These recommendations include expanding the test asset set beyond the standard 25 size/BM portfolios, imposing theory-driven constraints, employing GLS instead of OLS, and reporting confidence intervals. Our proposed method closely aligns with these recommendations, effectively addressing existing gaps in the context of beta-pricing models.

For Bayesian methods in asset pricing model comparison, Avramov and Chao (2006) propose a conditional model. Barillas and Shanken (2018) develop a Bayesian counterpart to the GRS test for traded factor models, employing Jeffreys priors for betas and residual covariances, along with a Sharpe ratio bound prior on alpha. Chib, Zeng, and Zhao (2020) demonstrate that standard Jeffreys priors on model-specific nuisance parameters lead to invalid Bayes factors. They introduce a novel class of improper priors via invertible mappings (Chib and Zeng, 2020), ensuring valid marginal likelihoods and robust model comparisons. A detailed application of this framework is provided in Chib, Zhao, and Zhou (2024). However, until the framework developed here, Bayesian model comparison methods have primarily focused on traded factors in time-series regressions. The approach of Allena (2020) considers nontraded factors but does not model them directly, relying instead on mimicking portfolios that introduce approximation error.

This paper also contributes to the literature on the inference of factor risk premia. Estimates of risk premia based on two-pass regression are often compromised by the presence of weak factors (e.g., Kan and Zhang, 1999; Kleibergen, 2009). To address this, Gospodinov, Kan, and Robotti (2019) conduct inference in the beta-pricing model using MLE under potential model misspecification. Building on these ideas, Giglio and Xiu (2021) demonstrate how omitting priced factors can bias standard risk premia estimators and propose a three-pass method to address this issue. Furthermore, Giglio,

Xiu, and Zhang (2024) introduce a supervised principal component analysis approach to recover risk premia associated with weak factors. While these frameworks represent notable progress, they do not provide a unified approach for simultaneous model comparison and risk premia estimation. This paper fills that gap by systematically comparing competing models and using the one that best explains the data for risk premia estimation. Models containing weak or redundant factors are discarded in the process, ensuring that inference is based on the best performing specification.

The paper is structured as follows. Section 2 describes the methodology. Section 3 provides simulation results. Section 4 outlines key empirical findings. We revisit the model comparison in Kan et al. (2013) (Section 5) and consumption-based model comparison in Kleibergen and Zhan (2020) (Section 6). Finally, Section 7 concludes.

2 Methodology

Traditional methods for estimating and testing the beta-pricing model include two-pass regression, MLE, and GMM. The estimation of risk premia depends critically on the specified model, with different models producing varying outcomes. While the "true" model remains unknown, some models demonstrably outperform others. Consequently, it is meaningful to focus on the risk premia derived from the best-fitting model within a given set of factors and test assets.

This section introduces the Bayesian counterpart to the beta-pricing model. Section 2.1 outlines the baseline model setup, while Section 2.2 discusses alternative estimation and comparison methods. Section 2.3 specifies the prior for the baseline model. The MCMC estimation procedure is detailed in Section 2.4, and Section 2.5 describes marginal likelihood calculations. In Section 2.6, we propose a method to assess the weakness of individual factors. Finally, Section 2.7 focuses on identifying the best-fitting model from a set of factors.

2.1 Model Set Up

We begin with presenting the notations and assumptions for estimating the linear factor model in the beta-pricing form. Let $\boldsymbol{r}_t : n \times 1$ denote the vector of test asset

excess returns and suppose that

$$\boldsymbol{f}_t = \begin{pmatrix} \boldsymbol{f}_t^T \\ \boldsymbol{f}_t^{NT} \end{pmatrix} : k = (k^T + k^{NT}) \times 1$$

is a vector of k assumed risk factors at time t, where we use superscript ^{T} to denote traded factors (the market and zero-investment returns) and ^{NT} for nontraded factors (innovations in macroeconomic variables) that affect excess returns of traded assets. Then, the linear factor model in the time series (TS) takes the form

$$\begin{aligned} \boldsymbol{r}_t &= \boldsymbol{a} + \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{e}_t \\ &= \boldsymbol{a} + \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{B}^{NT}\boldsymbol{f}_t^{NT} + \boldsymbol{e}_t \end{aligned}$$

where $\boldsymbol{B} = [\boldsymbol{B}^T \ \boldsymbol{B}^{NT}]$ and $\boldsymbol{e}_t$ is a vector of time-series random errors at time t. In the cross section (CS), we have the restrictions[2]

$$\begin{aligned} E(\boldsymbol{r}_t) &= \boldsymbol{B}\boldsymbol{\lambda} \\ &= \boldsymbol{B}^T\boldsymbol{\lambda}^T + \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT}\ , \ t \leq T \end{aligned}$$

where we have used a similar partition, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}^{T\prime}, \boldsymbol{\lambda}^{NT\prime})'$. In addition, we have the following restriction on the risk premia of traded risk factors:

$$\boldsymbol{\lambda}^T = E[\boldsymbol{f}_t^T]\,.$$

The best way to proceed is to integrate these two restrictions into the TS model. Taking expectations in the TS model, we get that

$$\boldsymbol{a} = \boldsymbol{B}^{NT}(\boldsymbol{\lambda}^{NT} - E[\boldsymbol{f}_t^{NT}]).$$

Now let $\boldsymbol{\mu}^{NT} = E[\boldsymbol{f}_t^{NT}]$. Then, the TS model that integrates both the CS restrictions and the traded risk factor constraints takes the form:

$$\boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t. \tag{1}$$

If there is only one NT factor, this restricted time series version of the beta-pricing model reduces to the formulation in Pastor and Stambaugh (2003).

If one supposes that $\boldsymbol{\mu}^{NT}$ is unknown then the restricted TS model can be supple-

[2]We exclude an intercept (zero-beta rate) as we use excess returns, assuming the zero-beta rate equals the risk-free rate. The model with a zero-beta rate is derived in Appendix A.VI.

mented with the model

$$\boldsymbol{f}_t^{NT} = \boldsymbol{\mu}^{NT} + u_t, \tag{2}$$

where the errors are uncorrelated and estimated jointly with the restricted TS model. This is the approach of Pastor and Stambaugh (2003) where the estimation of parameters $\boldsymbol{\theta} = [\boldsymbol{\lambda}^{NT}, \boldsymbol{B}^{T}, \boldsymbol{B}^{NT}, \boldsymbol{\mu}^{NT}]$ is achieved by GMM (their equations 17 and 18) by minimizing $\mathrm{g}(\boldsymbol{\theta})'\boldsymbol{W}\mathrm{g}(\boldsymbol{\theta})$, where $\mathrm{g}(\boldsymbol{\theta}) = (1/T)\sum_{t=1}^{T}\mathrm{f}_t(\boldsymbol{\theta})$,

$$\mathrm{f}_t(\boldsymbol{\theta}) = \begin{pmatrix} \mathrm{h}_t \otimes \boldsymbol{e}_t \\ \boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT} \end{pmatrix}, \quad \mathrm{h}_t' = \begin{pmatrix} 1 & \boldsymbol{f}_t^{T\prime} & \boldsymbol{f}_t^{NT\prime} \end{pmatrix},$$

$$\boldsymbol{e}_t = \boldsymbol{r}_t - \boldsymbol{B}^{NT}\left(\boldsymbol{\lambda}^{NT} - \boldsymbol{\mu}^{NT}\right) - \boldsymbol{B}^{T}\boldsymbol{f}_t^{T} - \boldsymbol{B}^{NT}\boldsymbol{f}_t^{NT}, \tag{3}$$

and $\boldsymbol{W}$ is the consistent estimator of the optimal weighting matrix.

Now, consider comparing alternative beta-pricing models, particularly those using different combinations of T and NT factors within the restricted TS framework. A key challenge arises because the dimensionality of the dependent variables in Equations (1) and (2) varies across models. As a result, direct comparison using marginal likelihoods is not feasible. Specifically, including only the NT factors relevant to each return model leads to differences in the response variables across models.

Therefore, to enable valid comparisons across different factor models, we propose an innovation: augmenting the return model with a model that includes the full set of factors. Specifically, the joint model we estimate is given by:

$$\begin{aligned} &\text{RETURN MODEL} && \boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^{T}\boldsymbol{f}_t^{T} + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t \\ &\text{FACTOR MODEL} && \boldsymbol{f}_t^{*} = \boldsymbol{\mu} + \boldsymbol{u}_t \end{aligned} \tag{4}$$

where $\boldsymbol{f}_t^{*} = (\boldsymbol{f}_{t1}, \boldsymbol{f}_{t2}) : (k_1^{*} + k_2^{*}) \times 1$, $\boldsymbol{f}_{t1}$ is the full set of traded factors, $\boldsymbol{f}_{t2}$ is the full set of nontraded factors, and $\boldsymbol{f}_t^{T}$ and $\boldsymbol{f}_t^{NT}$ are the subset of traded and nontraded factors, respectively, in the RETURN MODEL, and the first k^{T} elements of $\boldsymbol{\mu}$ in the FACTOR MODEL equal $\boldsymbol{\lambda}^{T}$.

One can interpret this joint model as a hierarchical model. The complete factor set is generated via the FACTOR MODEL, while returns are derived using the RETURN MODEL based on some subset of $\boldsymbol{f}_t^{*}$. These models are interlinked through the param-

eter $\boldsymbol{\mu}^{NT}$, representing a subset of the last k_2^* elements of $\boldsymbol{\mu}$. Together, they facilitate parameter estimation and guide the selection of the most relevant factors from $\boldsymbol{f}_t^*$ for explaining returns.

We complete this joint model by supposing that the innovations follow

$$\boldsymbol{e}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_e), \quad \boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$$

where $\boldsymbol{\Sigma}_e = \text{diag}(\sigma_1^2, ..., \sigma_n^2)$, and $\boldsymbol{\Sigma}_u$ is a full covariance matrix, respectively. The diagonality assumption on $\boldsymbol{\Sigma}_e$ allows for the fitting of large cross-sectional dimensions. It is not a limiting assumption since marginalized over the factors, the distribution of the returns has a full covariance structure.

Special Cases Notably, our model reduces to special cases when we consider only one type of factor at a time. With only NT factors, our model simplifies to:

$$\boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t. \tag{5}$$

which corresponds to Equation (14.209) in Cochrane (2009).

When there are only traded factors, we get:

$$\boldsymbol{r}_t = \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{e}_t. \tag{6}$$

which corresponds to Equation (14.202) in Cochrane (2009).

2.2 Alternative Estimation Methods

In this section, we review several traditional standard methods to estimate the beta-pricing model, but mostly overlook the traded factor constraint. Let $\boldsymbol{R} = [\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_T]'$, in the traditional two-pass approach, a CS regression is run of the sample mean of return $\bar{\boldsymbol{R}}$ on $\widehat{\boldsymbol{B}}$ to estimate $\boldsymbol{\lambda}$ in the second pass (Kan et al., 2013). Given the weighting matrix $\boldsymbol{W}$, the $\boldsymbol{\lambda}$ can be estimated as:

$$\widehat{\boldsymbol{\lambda}} = (\widehat{\boldsymbol{B}}'\boldsymbol{W}\widehat{\boldsymbol{B}})^{-1}\widehat{\boldsymbol{B}}'\boldsymbol{W}\bar{\boldsymbol{R}}.$$

For OLS, the weighting matrix $\boldsymbol{W}$ is the identity matrix. In GLS, $\boldsymbol{W}$ is set as the inverse of the residual covariance matrix in the initial pass. However, estimating the inverse covariance matrix can be computationally challenging when the number of test assets

is large. A practical alternative is the WLS estimator, where $\boldsymbol{W}$ is the inverse of the diagonal elements of the residual covariance matrix.[3] In simulations and empirical tests, when the test asset count is large, we opt for WLS over GLS.

Beyond two-pass regression, another approach to estimating the beta-pricing model is MLE. To derive MLE, we combine the TS and CS models into a single likelihood. The CS regression can be viewed as a restriction on the TS regression. By imposing the pricing restriction in the CS model and taking expectations within the time-series model, then substituting in the TS model, we get the restricted TS model (Cochrane, 2009, Equation 14.209),

$$\boldsymbol{r}_t = \boldsymbol{B}\boldsymbol{\lambda} + \boldsymbol{B}(\boldsymbol{f}_t - E[\boldsymbol{f}_t]) + \boldsymbol{e}_t. \tag{7}$$

For MLE, an auxiliary statistical model in the errors and factors that are i.i.d. normal and uncorrelated with each other, is introduced:

$$\boldsymbol{f}_t = E(\boldsymbol{f}_t) + \boldsymbol{u}_t,$$

$$\begin{bmatrix} \boldsymbol{e}_t \\ \boldsymbol{u}_t \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \boldsymbol{\Sigma} & 0 \\ 0 & \boldsymbol{V} \end{bmatrix}\right).$$

Then, the likelihood function is given by

$$\mathcal{L} \propto -\frac{1}{2}\sum_{t=1}^{T} \boldsymbol{e}_t'\boldsymbol{\Sigma}^{-1}\boldsymbol{e}_t - \frac{1}{2}\sum_{t=1}^{T} \boldsymbol{u}_t'\boldsymbol{V}^{-1}\boldsymbol{u}_t.$$

A more general approach to estimate is GMM, which relaxes the independent and identical distribution assumption of return. Rewrite TS and CS models into moment conditions (Cochrane, 2009, equation 12.188):

$$\boldsymbol{g}_T = \begin{bmatrix} E\left(\boldsymbol{r}_t - \boldsymbol{a} - \boldsymbol{B}\boldsymbol{f}_t\right) \\ E\left[\left(\boldsymbol{r}_t - \boldsymbol{a} - \boldsymbol{B}\boldsymbol{f}_t\right)\boldsymbol{f}_t\right] \\ E\left(\boldsymbol{R} - \boldsymbol{B}\boldsymbol{\lambda}\right) \end{bmatrix}. \tag{8}$$

GMM solves the minimization of $\boldsymbol{g}_T\boldsymbol{W}\boldsymbol{g}_T$ for some weighting matrix. An explicit example of GMM estimation under the general model with both traded and nontraded factors and with traded factor constraint is given in equation 3.

[3] Some approaches employ the return covariance matrix rather than the residual covariance (e.g., Kan et al., 2013).

Model Comparison Although many methods have been proposed to estimate beta-pricing models, comparison methods remain rare. The most commonly used performance measure CSR^2 for model comparison is calculated as follows,

$$R^2 = 1 - \frac{\boldsymbol{e}'\boldsymbol{W}\boldsymbol{e}}{\boldsymbol{e}_0'\boldsymbol{W}\boldsymbol{e}_0}, \quad \boldsymbol{e} = \bar{\boldsymbol{R}} - \boldsymbol{B}\widehat{\boldsymbol{\lambda}}, \quad \boldsymbol{e}_0 = [\boldsymbol{I}_N - \boldsymbol{1}_N(\boldsymbol{1}_N'\boldsymbol{W}\boldsymbol{1}_N)^{-1}\boldsymbol{1}_N'\boldsymbol{W}]\bar{\boldsymbol{R}}$$

and the weighting matrix $\boldsymbol{W}$ depends on the estimation method.

Alongside CSR^2, we consider the H-J Distance (Hansen and Jagannathan, 1997; Kan and Robotti, 2009), a widely used metric. The computation utilizes the two-stage GMM-SDF framework. Let the covariance matrix between $\boldsymbol{R}$ and $\boldsymbol{F}$ be $\boldsymbol{C}$, and the demeaned factor $\boldsymbol{F}$ be $\tilde{\boldsymbol{F}}$. The SDF loading $\boldsymbol{b}$ is defined as $(\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\bar{\boldsymbol{R}}$. The estimated normalized SDF is given by $\boldsymbol{m} = 1 - \tilde{\boldsymbol{F}} * \widehat{\boldsymbol{b}}$, with residuals $\boldsymbol{Rm}$. The squared H-J distance is $E(\boldsymbol{Rm})'E(\boldsymbol{Rm})/N^2$, normalized by the number of test assets N.

However, both CSR^2 and the H-J distance suffer from the limitation that a "larger" model always performs at least as well as a "smaller" one. To address this, we employ a Bayesian framework. By leveraging marginal likelihood, the Bayesian approach circumvents these challenges. Notably, the model maximizing marginal likelihood has an appealing property: in large samples, it converges to the true model or the closest one in Kullback-Leibler distance (Chib et al., 2018).

2.3 Prior Specification

This section outlines the prior specification for the baseline model, incorporating both traded and nontraded factors. To enable valid model comparisons, the model-specific prior must be proper, integrating to one over the parameter space. Moreover, prior distributions should remain comparable across models to ensure marginal likelihood differences are not driven by prior variations. Importantly, these priors should require minimal user input. Below, we detail the specification of a model-specific prior. The parameters of a restricted factor model are defined as follows:

$$\boldsymbol{\theta} = (\boldsymbol{B}, \boldsymbol{\lambda}^{NT}, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_u, \boldsymbol{\mu}), \quad \boldsymbol{B} = (\boldsymbol{\beta}_1', ..., \boldsymbol{\beta}_n')'. \tag{9}$$

We construct our model-specific prior by utilizing an initial portion of the data, referred to as the training sample, to calibrate the hyperparameters of each prior dis-

tribution. In our primary application, this training sample spans the first 10 years starting from January 1985, encompassing $T_0 = 120$ observations in total, denoted as $[\boldsymbol{R}_0, \boldsymbol{F}_0]$, where $\boldsymbol{F}_0 = [\boldsymbol{F}_0^T, \boldsymbol{F}_0^{NT}]$. The training sample primarily serves to determine the mean of the prior distribution. Conversely, the dispersion of the prior distribution is predominantly governed by user-specified hyperparameters, ensuring sufficient uncertainty regarding its central location. Notably, the training sample is excluded from both the estimation and model comparison stages.

Our prior for $\boldsymbol{\mu}$ is given by:

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{M}_0),$$

where

$$\boldsymbol{\mu}_0 = \bar{\boldsymbol{F}}_0, \quad \boldsymbol{M}_0 = \frac{4}{T_0}(\boldsymbol{F}_0 - \mathbf{1}_{T_0}\boldsymbol{\mu}_0')'(\boldsymbol{F}_0 - \mathbf{1}_{T_0}\boldsymbol{\mu}_0').$$

The factor of 4 in this expression is a user-specified hyperparameter to widen the distribution around the training sample mean.

For each column of $\boldsymbol{B}$ we suppose that

$$\boldsymbol{\beta}_i \sim \mathcal{N}(\boldsymbol{\beta}_{0i}, \boldsymbol{B}_{0i}),$$

where

$$\boldsymbol{\beta}_{0i} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{r}_{0i}, \quad \boldsymbol{X} = [\mathbf{1}_{T_0}, \boldsymbol{F}_0]\ ,$$

$\boldsymbol{r}_{0i}$ is the ith column of $\boldsymbol{R}_0$,

$$\boldsymbol{B}_{0i} = 4\sigma_{0i}^2(\boldsymbol{F}_0'\boldsymbol{F}_0)^{-1},$$

and σ_{0i}^2 is the residual variance in the time series regression of $\boldsymbol{r}_{0i}$ on $\boldsymbol{F}_0$.

Prior for risk premia of nontraded factors is obtained by taking the prior mean $\boldsymbol{\mu}_0$ and prior mean $\boldsymbol{B}_0$ and fitting the model $\boldsymbol{r}_t - \boldsymbol{B}_0^T\boldsymbol{f}_t^T - \boldsymbol{B}_0^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}_0^{NT}) = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{e}_t$ on the training sample data to give

$$\boldsymbol{\lambda}^{NT} \sim \mathcal{N}(\boldsymbol{\lambda}_0^{NT}, \boldsymbol{L}_0^{NT}).$$

Next, the prior of σ_i are

$$\sigma_i \sim \mathcal{IG}\left(\frac{\nu_{e,0i}}{2}, \frac{\delta_{e,0i}}{2}\right),$$

where

$$\nu_{e,0i} = 2(e_i^2/s_i^2 + 2), \quad \delta_{e,0i} = 2e_i(e_i^2/s_i^2 + 1),$$

and e_i is i^{th} residual from the regression $\boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}_0^{NT} + \boldsymbol{B}_0^T\boldsymbol{f}_t^T - \boldsymbol{B}_0^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}_0^{NT}) + \boldsymbol{e}_t$ on the training sample data and s_i^2 is the estimated ith residual variance from this regression. Finally, the prior for $\boldsymbol{\Sigma}_u$ is

$$\boldsymbol{\Sigma}_u \sim \mathcal{IW}(\rho_{u,0}, \boldsymbol{Q}_{u,0}),$$

where

$$\boldsymbol{Q}_{u,0} = (\boldsymbol{F} - \mathbf{1}_{T_0}\boldsymbol{\mu}_0')'(\boldsymbol{F} - \mathbf{1}_{T_0}\boldsymbol{\mu}_0'), \quad \rho_{u,0} = T_0.$$

In summary, the training sample prior employs a sequential Bayesian learning approach to automatically set priors. It reflects beliefs about parameter values as of December 1994, just before the estimation period. This method centers the prior distribution while adjusting variances and covariances upward to account for uncertainty. Importantly, the prior adapts to the model, ensuring fairness and comparability across specifications that would otherwise be challenging to achieve. Requiring no user intervention, this approach facilitates the estimation and comparison of numerous models with minimal effort.

2.4 MCMC Simulation Algorithm

We now proceed to sample the posterior distribution using Markov Chain Monte Carlo (MCMC) methods, with a focus on the general framework that accommodates both traded and nontraded factors. Detailed discussions of the two special cases can be found in Appendices A.IV and A.V. The posterior distribution of the model parameters, conditional on the observed data for $t = 1, \ldots, T$, is proportional to the product of the likelihood function and the prior distribution. The likelihood for the return

model is given by:

$$p(\boldsymbol{r}_t \mid \boldsymbol{f}_t^*, \boldsymbol{B}^T, \boldsymbol{B}^{NT}, \boldsymbol{\lambda}^{NT}, \boldsymbol{\Sigma}_e) = \prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{r}_t \mid \boldsymbol{B}^T \boldsymbol{f}_t^T + \boldsymbol{B}^{NT} \boldsymbol{\lambda}^{NT} + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}), \boldsymbol{\Sigma}_e\right),$$

where $\boldsymbol{\Sigma}_e = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, and the factor model likelihood is

$$p(\boldsymbol{f}_t^* \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}_u) = \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{f}_t^* \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}_u).$$

The posterior distribution is proportional to the product of the two likelihoods and the prior on the parameters:

$$p(\boldsymbol{B}^T, \boldsymbol{B}^{NT}, \boldsymbol{\lambda}^{NT}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_u).$$

We employ MCMC methods to sample from the resulting posterior distribution. Specifically, we sequentially sample from the conditional distributions of parameter blocks, given the remaining parameters, to construct a Markov chain whose invariant distribution corresponds to the posterior distribution (Chib, 2001). The conditional distributions for these parameter blocks are derived based on the following formulations of the joint model. For updating $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_e$ we write the joint model as

$$\boldsymbol{r}_t = \boldsymbol{B}\tilde{\boldsymbol{f}}_t + \boldsymbol{e}_t,$$

where

$$\tilde{\boldsymbol{f}}_t = \begin{pmatrix} \boldsymbol{f}_t^T \\ \boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT} + \boldsymbol{\lambda}^{NT} \end{pmatrix}, \tag{10}$$

from which the conditional posterior distributions of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_e$ can be derived by Bayesian results for multivariate regression models. The updating of $\boldsymbol{\Sigma}_u$ arises easily from the model

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t. \tag{11}$$

For updating $\boldsymbol{\lambda}^{NT}$ we rewrite the full model now as

$$\boldsymbol{r}_t - \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{e}_t.$$

Finally, for updating $\boldsymbol{\mu}$ we write the joint model as

$$\boldsymbol{r}_t - \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} - \boldsymbol{B}\boldsymbol{f}_t = -\boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} + \boldsymbol{e}_t, \tag{12}$$

and

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t,$$

and then derive the conditional posterior of $\boldsymbol{\mu}$ in two steps, first with the return model likelihood and then with the factor model likelihood. We supply the full details in Appendix A.III.

Thus far, we have developed a model incorporating both traded and nontraded factors. Appendix A.IV details the MCMC algorithm for the specification with only nontraded factors (e.g., CCAPM in Breeden et al. (1989)). The procedure is similar but excludes the traded factor component. Appendix A.V provides the MCMC algorithm for the specification with only traded factors (e.g., CAPM). This estimation is streamlined as the parameter $\boldsymbol{\lambda}^{NT}$ is omitted. While focusing solely on traded factors, this approach deviates from the standard single-factor time-series model by imposing cross-sectional constraints and explicitly modeling factor covariance structures.

2.5 Marginal Likelihood

Our primary objective in this empirical study is not solely to estimate a single model but also to conduct a comprehensive comparison of multiple competing models. In the Bayesian framework, alternative models are evaluated using posterior odds based on the observed data. The posterior odds provide a quantitative measure of the relative support that the data offer for each competing model. Let $p_j = \Pr(\mathcal{M}_j)$ denote the prior probability assigned to model $\mathcal{M}_j$. The posterior odds between model $\mathcal{M}_i$ and $\mathcal{M}_k$ are

$$\frac{\Pr(\mathcal{M}_i \mid \boldsymbol{y}_{1:T})}{\Pr(\mathcal{M}_k \mid \boldsymbol{y}_{1:T})} = \frac{p_i}{p_k} \frac{m(\boldsymbol{y}_{1:T} \mid \mathcal{M}_i)}{m(\boldsymbol{y}_{1:T} \mid \mathcal{M}_k)},$$

where the marginal likelihood is defined as the integral of the likelihood function with respect to the prior density

$$m(\boldsymbol{y}_{1:T} \mid \mathcal{M}_j) = \int_{\Theta_j} p(\boldsymbol{y}_{1:T} \mid \boldsymbol{\theta}, \mathcal{M}_j) \pi(\boldsymbol{\theta} \mid \mathcal{M}_j) \mathrm{d}\boldsymbol{\theta}.$$

When prior odds across models are set to one, as assumed here, posterior odds reduce to the ratio of marginal likelihoods. Thus, marginal likelihood becomes the key

metric for model performance. By inherently penalizing model complexity, marginal likelihoods prevent more flexible models from ranking higher merely due to overfitting noise. Moreover, model selection based on Bayes factors or marginal likelihoods exhibits desirable asymptotic properties. If the true model is among the candidates, the model with the highest marginal likelihood is selected with probability approaching one as the sample size increases. If the true model is absent (a more realistic scenario), the model with the highest marginal likelihood approximates the true model in terms of Kullback–Leibler distance (Chib et al., 2018). Details on marginal likelihood computation are provided in Appendix A.III.

2.6 Absolute Comparison of One (Weak) Nontraded Factor

When introducing a nontraded factor, the first step is to assess whether it is weak or useless. If the model is identified, i.e., the $\boldsymbol{\beta}$ matrix is full rank, the two-pass procedure provides consistent and asymptotically normal risk premia estimates, enabling standard hypothesis testing (e.g., t-statistics) in a frequentist framework. However, with a weak or useless factor (i.e., $\boldsymbol{\beta}_i = 0$), identification fails, distorting parameter inference. Our Bayesian approach also estimates $\boldsymbol{\beta}$, but integrates it out numerically when computing the marginal likelihood. This ensures valid model comparisons even with weak factors. While weak factors may bias $\boldsymbol{\beta}$ estimates and risk premia, the model's marginal likelihood declines relative to specifications excluding such factors.

An alternative definition of a weak factor arises when $\boldsymbol{\beta}$ lacks sufficient cross-sectional variation, a concern particularly relevant for estimating the zero-beta rate. Including an intercept in the cross-sectional regression augments the factor loading matrix to $(\mathbf{1}_N : \boldsymbol{\beta})$. In this case, a column of identical β values reduces the matrix's rank. By contrast, under our framework, which excludes the zero-beta rate, risk premia remain well-identified, allowing consistent β values.

In the current framework, $\boldsymbol{f}^T$ denotes the set of traded factors included in the return-generating model, while $\boldsymbol{f}^{NT}$ represents the nontraded factors. Specifically, we examine a single nontraded factor alongside a benchmark model that includes only traded factors, such as the Fama-French three-factor (FF3) model. The model based

solely on traded factors, $\boldsymbol{f}^T$, is referred to as M^T, whereas the model that integrates both traded and nontraded factors, $(\boldsymbol{f}^T, \boldsymbol{f}^{NT})$, is denoted by M.

Result. The nontraded factor is weak for specific test assets if the marginal likelihood of M^T exceeds that of M given the benchmark factors. In essence, if introducing a new nontraded factor does not enhance the marginal likelihood, it likely exhibits a weak correlation with test asset returns, even if its risk premia estimate is nonzero. Marginal likelihood inherently penalizes such weak factors, even when their estimated risk premia appear significant. As shown through simulations in Section 3.3, weak factors lead to lower marginal likelihoods for models including them, even when the true λ in the data-generating process (DGP) is large. Conversely, for non-weak factors—where $\boldsymbol{\beta}$ sufficiently identifies risk premia—model M achieves a higher marginal likelihood, even if the risk premia is small. However, in such cases, the risk premia estimate in M is often insignificant.

For instance, consider the FF3 model (MKT, SMB, HML) as the benchmark traded-factor framework. The nontraded factor of interest is the intermediary capital risk factor (HKMcapital). Using the traded-factor-only model, we compute the marginal likelihood, ML^T, based on the reduced specification (Proposition A.V). Including HKMcapital, we calculate the marginal likelihood, ML, from the full model (Proposition A.III). If $ML > ML^T$, we infer that HKMcapital adds explanatory power in pricing the cross-section of test assets. Next, we estimate the risk premia to evaluate the significance of HKMcapital and assess its economic relevance. If its contribution proves weak or risk premia are insignificant, the evidence may reject HKMcapital as a meaningful factor.

2.7 Simultaneous Comparison of A Set of Factors

In the previous section, we tested a single nontraded factor using a benchmark model. Here, we extend the analysis to compare multiple asset pricing models, encompassing both nested and non-nested structures with traded and nontraded factors. A key challenge is recognizing that none of the models may fully capture the true pricing

mechanism. While we impose the cross-sectional pricing constraint $\alpha = 0$, we do not assume any model strictly satisfies it. Instead, we aim to identify the model that best approximates this constraint. To this end, we interpret marginal likelihoods and posterior probabilities as indicators of relative model performance. Marginal likelihood assesses how well pricing constraints are met and how effectively factors explain time-series variations. These metrics integrate cross-sectional and time-series information, offering a comprehensive evaluation of model fit.

As outlined earlier, our method evaluates the pricing performance of a set of factors on test assets by imposing a pricing constraint. To ensure comparability of marginal likelihoods, introducing $\boldsymbol{f}^*$ (the full factor set) simplifies model comparisons. The key distinction lies in whether the model includes only traded factors, only nontraded factors, or both. For models with only traded or nontraded factors, the reduced model should be used to compute marginal likelihoods, as detailed in Appendices A.IV and A.V. Models combining both factor types should rely on the full model.

This setup simplifies the comparison of factor sets while addressing the weak factor issue. Models with weak factors are penalized in marginal likelihoods, as such factors fail to explain both time-series variation and cross-sectional pricing restrictions effectively. As discussed, models with weak factors yield lower marginal likelihoods. Factors excluded from the model with the highest marginal likelihood lack additional explanatory power. Thus, risk premia estimation should center on the top model, as estimates from alternative models may be unreliable or spurious.

Model comparison for nontraded factors using CSR^2 or related metrics is well-documented (Kan et al., 2013). Our approach reexamines these comparisons through a Bayesian lens. Models with negligible CSR^2 improvements are unlikely to yield higher marginal likelihoods. In Section 5, we first assess the eight predefined models from the original study to identify the best performer. We then broaden the analysis by evaluating all possible factor combinations and exploring potential enhancements over the original top model.

Among nontraded factors, consumption growth is central, with strong theoretical

underpinnings but difficult to measure. To address this, the literature has proposed various proxies for consumption growth risk. These efforts underscore the importance of robust estimation techniques, making our approach well-suited for such analysis. By incorporating multiple consumption measures alongside the MKTRF factor as a traded control, our method evaluates the relative efficacy of different proxies. This is achieved by comparing these measures within a unified framework or using the MKTRF factor as a test asset, as detailed in Section 6.

3 Simulation

This section assesses the finite-sample performance of our Bayesian estimators via Monte Carlo simulations. We conduct four experiments: Section 3.1 investigates risk premia estimation under the correctly specified model. Section 3.2 evaluates the effect of the traded factor constraint. Section 3.3 examines nontraded factors under weak identification. Lastly, Section 3.4 compares model performance across factor sets. All parameters are calibrated using the true data described in Section 4.1.

3.1 Simulation under Correctly Specified Model

The primary focus of the beta-pricing model is the estimation of risk premia. Before coming into the model comparison part, we examine whether our model can give comparable estimation results under the correctly specified true model with traditional approaches. We first simulate factors from:

$$\boldsymbol{f}_t = \boldsymbol{\mu} + \boldsymbol{u}_t$$

The $\boldsymbol{\mu}$ and the covariance of factors $\boldsymbol{\Sigma}_u$ are calibrated using the true factor data. We then simulate the asset returns from the following model:

$$\boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^{T}\boldsymbol{f}_t^{T} + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t$$

where $\boldsymbol{r}_t$ contains 100 Fama-French sorted portfolios, $\boldsymbol{f}^T$ contains 6 commonly used traded factors MKTRF, SMB, HML, RMW, CMA, and UMD, while $\boldsymbol{f}^{NT}$ contains two nontraded factors: durable consumption expenditure (PCEDG) and intermediary capital (HKMcapital). The factor loadings $B = [\boldsymbol{B}^T \ \boldsymbol{B}^{NT}]$ and the residual covariance $\boldsymbol{\Sigma}_e$

are calibrated from OLS regression of portfolios on these 8 factors with intercept. The risk premia of nontraded factors $\boldsymbol{\lambda}^{NT}$ is set as $(0.002, 0.003)$, comparable with that of traded factors. The sample period T exactly matches the empirical data.

Table 1: **Simulation Evidence: Estimation under Correctly Specified Model**

This table presents simulation results for the Correctly Specified model. The analysis incorporates 8 factors —– 6 traded and 2 nontraded. Parameters are calibrated using true test assets and factors. The table reports bias and RMSE for WLS and Bayesian estimation, both with and without the traded factor constraint (NC). Results are based on 1,000 simulations and expressed in percentages.

	WLS bias	Bayes bias	BayesNC bias	WLS RMSE	Bayes RMSE	BayesNC RMSE
			Traded Factors			
MKTRF	-0.001	-0.004	0.177	0.209	0.213	0.277
SMB	0.004	0.000	0.041	0.139	0.141	0.148
HML	0.006	-0.010	-0.037	0.143	0.145	0.154
RMW	-0.010	-0.004	-0.042	0.119	0.118	0.128
CMA	0.000	0.001	-0.034	0.100	0.096	0.109
UMD	-0.029	0.002	-0.056	0.231	0.223	0.244
			Nontraded Factors			
PCEDG	-0.065	-0.062	-0.047	0.224	0.226	0.227
HKMcapital	0.050	0.024	0.221	0.342	0.345	0.412

Table 1 presents the bias and RMSE from 1,000 simulations. Table A.2 in the appendix reports the rejection frequency of the null hypothesis $H_0 : \boldsymbol{\lambda} = \boldsymbol{\lambda}_{true}$. The column "BayesNC" refers to our Bayesian estimation without the traded factor constraint. For all factors, our constrained Bayes estimate achieves bias and RMSE comparable to WLS. However, without the constraint, the model deviates from theory-implied values for traded factors, increasing bias and RMSE, particularly for these factors. This underscores the critical role of the traded factor constraint. Nontraded factors are also affected; for instance, in the case of HKMcapital, BayesNC exhibits significantly higher bias.

3.2 Simulation for Useless Factor under Traded Factor Constriant

We employ a simple simulation with traded factors to demonstrate the critical role of imposing restrictions on their risk premia estimation. Lewellen et al. (2010) use

a simulation of artificial factors with zero means to show that without this constraint, the cross-sectional regression just searches for the best possible fit in the cross-sectional regression and thus, the simulated CSR^2 differs from zero significantly, while in fact, it should be zero. Align with their method but slightly different, our simulation uses the marginal likelihood to show the importance of imposing this constraint.

Just as Lewellen et al. (2010)'s simulation study, we use data from empirical directly. Specifically, we use the same 5×5 size and book-to-market sorted portfolio returns and FF3 factors. We consider two kinds of useless factors. $ul1$ standard for a zero mean noisy factor, while $ul2$ also has zero mean, but correlated with return:

$$ul1_t = e_t, e_t \sim \mathcal{N}(0, (1\%)^2)$$

$$ul2_t = 0.5 * (\bar{r}_t - \bar{\bar{r}}) + e_t, e_t \sim \mathcal{N}(0, (1\%)^2)$$

where $\bar{r}_t$ is the cross-sectional average of test asset returns. Under this setup, these two factors should not have any pricing ability relative to other true factors. Under the imposed constraint, it is anticipated that the model incorporating this factor will underperform relative to the unconstrained model. In the absence of this constraint, the model optimizes for the best fit by allowing for a nonzero risk premium deviation, thereby enhancing the marginal likelihood.

Table 2 shows the results for this simulation. The number is the average of 100 simulation runs. Each row represents the useless factor, which is $ul1$ or $ul2$. Panel A uses the estimation without traded factor constraint, while Panel B imposes that constraint. Log ML reduced represents the log marginal likelihood without the chosen simulated factor, while log ML Full stands for the model with all 4 factors (FF3 factors plus one artificial factor). Panel C is the GLS CSR^2 and H-J distance for the reduced model, and Panel D for the full model.

In panel A, all factors exhibit increased marginal likelihood upon adding the corresponding simulated factor, despite its zero mean. In contrast, panel B shows a decline in marginal likelihoods when adding the redundant factor, suggesting no additional explanatory power. Comparing marginal likelihood across panels, it is expected

Table 2: **Evaluating Useless Factors under Traded Factor Constraint**

This table summarizes simulation results evaluating the impact of useless factors under the traded factor constraint. Panel A presents outcomes from the Bayesian model without the constraint, while Panel B incorporates it. The empirical dataset comprises 5×5 size and book-to-market sorted portfolios, with FF3 factors and one simulated zero-mean factor. ul_1 represents a zero-mean factor uncorrelated with returns, and ul_2 is zero-mean but correlated with returns. "Log ML Reduced" denotes the log marginal likelihood of the model excluding the simulated factor, while "Log ML Full" corresponds to the full model. Panel C reports GLS CSR^2 and H-J distance for the reduced model, and Panel D for the full model. Results are averaged over 100 simulations.

	log ML reduced	log ML Full	CSR^2(%)	HJD(%)
	Panel A: Bayes No Constraint		Panel C: GLS reduced	
$ul1$	29604.0	29660.6	19.78	0.0373
$ul2$	29605.0	29689.0	19.78	0.0373
	Panel B: Bayes With Constraint		Panel D: GLS Full	
$ul1$	29392.6	29373.5	23.45	0.0356
$ul2$	29377.2	29377.1	26.81	0.0350

to increase when the constraint is removed. However, this does not imply that the unconstrained model outperforms the constrained one. The unconstrained model disregards theoretical underpinnings and prioritizes fit, often deviating from the theoretical framework. Artificial factors should not enhance performance. Removing constraints can artificially inflate marginal likelihood, potentially leading to misguided conclusions. In addition, to verify the Lewellen et al. (2010)'s results, Panels C and D show that adding a useless zero-mean factor increases CSR^2 while decreasing the H-J distance, misleadingly suggesting the useless factor contributes to pricing.

3.3 Simulation with Weak Factors

Weak factor issues are prevalent in beta-pricing studies. In the presence of a weak factor, traditional two-pass methods yield invalid parameter inferences. Our approach mitigates this issue by excluding models with weak or useless factors through a comparison of marginal likelihoods. Models with weak factors exhibit lower marginal likelihoods, even if their risk premia appear statistically significant. Such significance, however, is unreliable as it fails to enhance pricing performance. Consequently, weak

factors are unlikely to be part of the correctly specified model, and estimating their risk premia is spurious.

We demonstrate in Section 3.1 that our estimate matches or surpasses WLS under the correct model specification. However, what if the factor is weak, particularly for nontraded factors? We conducted an additional simulation to test the model under a weak factor, where the beta is significantly smaller than the calibrated value. We still use the calibrated parameter set $[\boldsymbol{\mu}, \boldsymbol{\Sigma}_u, \boldsymbol{B}, \boldsymbol{\Sigma}_e]$. We derive returns from a 7-factor model comprising six traded factors calibrated to FF6 and a nontraded factor tied to the true nontraded factor. The sample period T aligns with the empirical data. We vary the beta scale using a parameter w.

$$\boldsymbol{r}_t = w\boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^T\boldsymbol{f}_t^T + w\boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t$$

To measure weakness, we multiply the w on the parameter of nontraded factors' calibrated beta in the corresponding simulation run. A smaller w means weaker; that is, beta is much smaller. We also consider different lambda to check how the weak factor and strength of risk premia jointly affect the marginal likelihood.

Table 3 contains four panels, each corresponding to a distinct risk premia level. Within each panel, five weakness levels for the nontraded factor are examined. A weakness level of 1 reflects the original calibrated beta, while a level of 0.01 reduces beta to 1% of its original value, making it 100 times smaller. The dML measures the difference in marginal likelihood between the FF6 model augmented with the nontraded factor and the baseline FF6 model. A positive dML indicates that the nontraded factor enhances pricing by improving marginal likelihood relative to the FF6 model.

The first two panels provide consistent insights: when the nontraded factor is sufficiently strong ($w \geq 1$), the model with HKMcapital achieves a higher marginal likelihood than the FF6-only model. Even with HKMcapital's λ as small as 0.0002 (relative to a market risk premium of 0.005), it outperforms. However, as the factor weakens, its marginal likelihood falls below that of the FF6 model, despite HKMcapital being part of the true data-generating process. Interestingly, equivalent levels of $wB^{NT}\lambda^{NT}$

Table 3: **Simulation Evidence: Evaluating Weak Factors**

The baseline model incorporates six traded factors. Each column augments the baseline with a distinct nontraded factor, while each panel corresponds to varying levels of risk premia, parameterized by weakness (w). Lower w indicates a smaller beta. dML denotes the log marginal likelihood difference between the traded-factor-only model and the model augmented with one nontraded factor. Parameters are calibrated using true test assets and factors, with results averaged over 100 simulations.

	dML	dML	dML	dML	dML	dML	dML	dML
	IndProd	Liq	LTY	M2_SA	PCEDG	PCEND	HKMcapital	PEAR
	Panel A: λ = 0.0002							
w=10	10847	8782	10471	8080	5576	12399	13862	5028
w=1	113	46	93	29	-27	150	200	-36
w=0.1	-113	-111	-109	-112	-111	-111	-109	-113
w=0.01	-113	-112	-110	-117	-111	-115	-111	-111
	Panel B: λ = 0.002							
w=10	11256	8670	10419	8959	5655	12630	13768	5033
w=1	120	47	94	49	-29	158	194	-36
w=0.1	-111	-112	-109	-114	-110	-111	-108	-112
w=0.01	-114	-111	-111	-114	-112	-112	-115	-112
	Panel C: λ = 0.02							
w=10	26729	9121	11178	37302	7316	27016	14087	5897
w=1	941	51	108	2034	-1	938	221	-24
w=0.1	-111	-110	-110	-106	-113	-109	-108	-111
w=0.01	-112	-112	-114	-114	-114	-112	-112	-112
	Panel D: λ = 0.2							
w=10	99189	32128	34763	117948	42851	98166	53409	30466
w=1	24271	1394	1696	37255	2739	23983	4887	1249
w=0.1	641	-107	-105	1829	-101	591	-79	-109
w=0.01	-114	-111	-112	-110	-112	-112	-111	-112

do not yield consistent results. For instance, when $w = 100$ and $\lambda^{NT} = 0.0002$, the marginal likelihood exceeds that of the FF6-only model. Conversely, when $w = 0.01$ and $\lambda^{NT} = 0.2$, it declines significantly, driven primarily by factor weakness.

3.4 Simulation for Model Comparison

Having shown that our method can exclude weak factors by comparing models with or without that factor, what's next is how to find the true (or the best among the full set) model given a set of candidate factors. In our framework, the selection of different sets of factors for the return model results in distinct factor models, which in

turn produce varying risk premia estimations for the same factor. Including redundant models is not only meaningless but also risks leading to incorrect inferences under potential misspecification. In this subsection, we conduct a simulation to compare models within a given set of factors.

We generate factors and returns similarly as section 3.1, but we generate more factors and only use one combination of factors to generate returns. Specifically, the full set of factors includes MKTRF, SMB, HML, RMW, CMA, and UMD. In contrast, the nontraded factors contain liquidity (Liq), presidential economic approval rating (PEAR), nondurable consumption expenditure (PCEND), and intermediary capital (HKMcapital). The true model used to generate a return is MKTRF, SMB, HML, plus HKMcapital. In this setting, there are $(2^6 - 1) * (2^4) = 1008$ model specifications. The factors that are not included can be considered weak or useless factors, meaning that $\boldsymbol{\beta}$ is equal to $\mathbf{0}$. The sample period T exactly matches the empirical data. We calculate the marginal likelihood of all these models, OLS and GLS CSR^2 and H-J distance.

We simulate 100 iterations, reporting the average log marginal likelihood, OLS GLS CSR^2, model rank by marginal likelihood, and H-J distance for each run. Additionally, we present the marginal likelihood and average rank without the traded factor constraint (denoted NC). Table 4 lists the top 10 models ranked by average marginal likelihood under the constraint. Only marginal likelihood with the constraint correctly identifies the true model as top-ranked. While CSR^2 increases with additional factors, the improvement is modest, making it unreliable for model selection. Consistent with CSR^2, adding factors reduces H-J distance. An average rank of 1 demonstrates that marginal likelihood reliably identifies the true model among nearly 1,000 candidates when strong factors are present across 100 simulations. Without the traded factor constraint, performance deteriorates sharply, with the true model achieving an average rank of only 20. In this case, marginal likelihood instead selects FF3 as the top model, highlighting the critical role of the traded factor constraint.

It is crucial to emphasize that not all traded factors need to be included in the model, as evidenced by the simulation results. Specifically, we generate six traded

Table 4: **Simulation Evidence: Model Comparison under Marginal Likelihood**

This table summarizes simulation results comparing 1008 model specifications with 10 factors: 6 traded and 4 non-traded. The true DGP is a four-factor model (MKTRF, SMB, HML, HKMcapital). Parameters are calibrated using the true test assets and factors. Metrics reported include log marginal likelihood, WLS CSR^2 (%), average model rank, and H-J distance (%). NC denotes the Bayesian model without traded factor constraints. Simulations are conducted 100 times.

Top 10 marginal likelihood models	logML	logMLNC	rank	rankNC	CSR^2	H-J D
MKTRF+SMB+HML+HKMcapital	55337.2	55475.0	1.0	20.5	62.6	0.0171
MKTRF+SMB+HML+CMA+HKMcapital	55284.7	55584.7	3.7	4.8	63.4	0.0168
MKTRF+SMB+HML+RMW+HKMcapital	55284.5	55595.1	3.8	3.2	63.3	0.0168
MKTRF+SMB+HML+UMD+HKMcapital	55284.1	55594.7	3.9	3.2	63.5	0.0168
MKTRF+SMB+HML+PEAR+HKMcapital	55278.8	55566.8	5.1	7.0	63.5	0.0168
MKTRF+SMB+HML+Liq+HKMcapital	55278.8	55595.3	5.1	3.1	63.6	0.0167
MKTRF+SMB+HML+PCEND+HKMcapital	55276.8	55572.4	5.6	6.2	63.4	0.0169
MKTRF+SMB+HML	55243.6	55759.2	11.5	1.0	48.0	0.0203
MKTRF+SMB+HML+RMW+CMA+HKMcapital	55231.2	55491.5	13.0	16.3	64.1	0.0165
MKTRF+SMB+HML+CMA+UMD+HKMcapital	55231.1	55485.0	13.1	19.4	64.2	0.0165

factors and set their risk premia equal to their respective factor means. However, if certain factors do not contribute to pricing, they are excluded from the model. This underscores an important insight: our method is capable of excluding not only weak nontraded factors but also weak traded factors. In this framework, the remaining three traded factors exhibit zero betas and can naturally be interpreted as weak traded factors, which are not selected by the top-performing model.

4 Empirical Results

This section summarizes our main findings. Section 4.1 outlines the dataset employed in our core analysis. Section 4.2 evaluates the empirical performance of two constraints on traded factors, as proposed by Lewellen et al. (2010). Section 4.3 analyzes the absolute weakness of individual nontraded factors. Section 4.4 reports model comparison results across all factor combinations. Lastly, Section 4.5 evaluate the model performance through tangency portfolio Sharpe Ratio.

4.1 Data

In our analysis, we consider a large cross section of test assets, that is, 302 equity portfolios from Ken French's website. It consists of 25 portfolios sorted by size and book-to-market ratio, 17 industry portfolios, 25 portfolios sorted by operating prof-

itability and investment, 25 portfolios sorted by size and variance, 35 portfolios sorted by size and net issuance, 25 portfolios sorted by size and accruals, 25 portfolios sorted by size and beta, 25 portfolios sorted by size and momentum, 25 portfolios sorted by size and operating profitability, 25 portfolios sorted by size and investment, 25 portfolios sorted by size and short-term reversal, and 25 portfolios sorted by size and long-term reversal. The sample ranges from January 1985 to December 2023.

We consider 8 prominent traded factors, including MKTRF, SMB, HML, CMA, RMW, UMD, BAB, and QMJ. The first six factors are downloaded from Ken French's website, while the last two factors are from the AQR website. For nontraded factors, we also collect 8 proposed factors in the literature, including IndProd, Liq, LTY, M2_SA, PCEDG, PCEND, HKMcapital, and PEAR.

For nontraded factors, it is typically the AR model innovations, rather than the factors themselves, that are of primary interest. For the factors Liq and HKMcapital, we directly download the data from the authors' websites. These factors are already treated as AR innovations in their original papers, so we use the downloaded series without further modifications. For the consumption growth factors, including PCEDG and PCEND, we calculate the real per capita growth rate using population and CPI data from FRED following the literature. For other factors, which are level factors obtained from sources like FRED or PEAR, we compute the AR(1) models and extract the innovations serving as the nontraded factors. All the definitions are in table A.1.

4.2 The Role of Traded Factor Constraint

As suggested by Lewellen et al. (2010), we consider a large cross section of test assets and impose the traded factor constraint. In this section, we show how these suggestions help improve the empirical test. We consider all subsets of test assets and 8 traded factors as described in section 4.1. To emphasize the role of traded factor constraint, we do not include nontraded factors in this section. The full set of factors and the subset of factors included in the return model are both these 8 traded factors. We estimate this factor model specification under different sets of test assets.

Table A.3 in the Appendix reports the estimation results of 8 traded factors and

different panel stands for different sets of test assets. Within each panel, we report three types of estimation: Bayes without traded factor constraint, Bayes with that constraint, and WLS (without constraint). For each factor, we report the posterior mean for Bayes (mean for WLS) and 5%/ 95% credible interval (confidence interval for WLS).

We show that for a traded factor, if we impose the constraint, its estimation is stable and remains almost unchanged across different test assets (as it is expected to be around the time-series mean). In contrast, without this constraint, the estimation varies significantly. For some traded factors, the problem is less severe, but for others, the significance may even change. For example, CMA has a highly insignificant estimate using almost all subsets of test assets. However, the constraint says that it should be slightly significant. BAB has a similar phenomenon, with an even more negative lower interval. On the contrary, the constraint suggests that it should be highly significant. Imposing the constraint makes the results more reasonable because the traded factor's risk premia, by theory, should be their mean.

To further emphasize the role of including more test assets, the last panel shows the estimation results using all 302 test portfolios. In this case, the situation has improved to some extent. For most factors, the risk premia estimates are much closer between the models with and without constraints compared to using only a subset of test assets, especially for CMA and BAB, which have a bad performance using a subset of test assets. More importantly, the posterior credible intervals are also close, reflecting a more accurate significance level.

In summary, these results confirm the two arguments in Lewellen et al. (2010) that the asset pricing tests can be improved by adding more test assets rather than 5×5 size and book-to-market sorted portfolios and also imposing theory-motivated constraints for traded factors.

4.3 Absolute Test of Single Nontraded Factor

Before evaluating all the factors of interest as a full set, we first assess the weakness of each nontraded factor while controlling for different sets of traded factors. Specifically, we analyze three benchmark traded factor models: CAPM, FF5, and All

traded factors (All). Panel A of table 5 presents the estimated risk premia and marginal likelihood values for each nontraded factor. For instance, the first row provides the estimated risk premium, while the second row shows the 95 % credible interval. The third row of "dML" indicates the difference in log marginal likelihood between two models with and without that nontraded factor. For example, for IndProd under MKTRF control, dML less than zero means that the model includes both MKTRF and IndProd under-performs the model with only MKTRF, while the full set of factors includes both MKTRF and IndProd. This comparison of marginal likelihood focuses on the absolute weakness of each nontraded factor, as discussed in the section 2.6. Panel B gives the estimation of WLS, and Panel C is the mimicking portfolio estimate with a 95% confidence interval in parentheses. We also report the mimicking portfolio constructed using all test assets for the nontraded factors.

Among all the nontraded factors, IndProd, Liq, M2_SA, and PCEND appear to be the weakest, as they are outperformed by the pure traded factor models across all benchmarks. While some of these factors show significant risk premia estimates, these results are likely spurious due to their inherent weaknesses. LTY only outperforms the simplest benchmark, CAPM, but fails against all other models. PCEDG and PEAR also outperform MKTRF but underperform in the remaining cases. Additionally, the risk premia estimates for these factors are highly sensitive to the choice of benchmark, further highlighting their weaknesses.

Unlike all other factors, HKMcapital demonstrates the strongest performance. It consistently exhibits a higher marginal likelihood across all cases compared to the traded factor benchmarks. The risk premium estimation for HKMcapital is consistent across all models and aligns closely with the WLS estimate. Additionally, the mimicking portfolio produces a very similar estimation and confidence interval. These results suggest that HKMcapital could be a potentially valuable nontraded factor.

4.4 Simultaneous Model Comparison

The previous section demonstrated that most of the nontraded factors do not provide additional pricing contributions. We now proceed to the simultaneous compar-

Table 5: **Evaluation of Single Nontraded Factors**

The table reports estimation results for individual nontraded factors, controlling for three benchmark traded factor models: CAPM, FF5, and All (FF6 + BAB + QMJ). Panel A presents Bayes estimation results, while Panel B reports WLS outcomes. dML denotes the difference in log marginal likelihood between the traded factor-only model and the traded factor plus one nontraded factor model. Mimicking portfolio estimates are shown in Panel C. 95% confidence (credible) intervals are in parentheses.

	IndProd	Liq	LTY	M2_SA	PCEDG	PCEND	HKMcapital	PEAR
				Panel A: Bayes				
CAPM	-0.50	3.77	0.74	-0.35	0.34	-0.57	1.68	-3.58
	[-0.67,-0.32]	[2.64,4.84]	[0.08,1.37]	[-0.42,-0.28]	[0.00,0.68]	[-0.73,-0.41]	[1.02,2.33]	[-4.27,-2.93]
	dML: -271	dML: -245	dML: 336	dML: -219	dML: 304	dML: -442	dML: 1213	dML: 38
FF5	0.00	-1.11	0.00	0.15	-1.16	-0.14	0.73	-0.69
	[-0.14,0.14]	[-2.08,-0.13]	[-0.87,0.89]	[0.08,0.22]	[-1.74,-0.55]	[-0.29,0.01]	[-0.02,1.48]	[-1.82,0.44]
	dML: -223	dML: -185	dML: -367	dML: -311	dML: -87	dML: -101	dML: 108	dML: -123
All	-0.10	-2.15	2.37	0.15	-0.04	-0.28	1.38	-0.23
	[-0.24,0.03]	[-3.23,-1.10]	[1.53,3.18]	[0.07,0.22]	[-0.94,0.93]	[-0.43,-0.13]	[0.59,2.19]	[-1.69,1.43]
	dML: -166	dML: -192	dML: -201	dML: -242	dML: -103	dML: -77	dML: 22	dML: -114
				Panel B: WLS				
CAPM	-0.11	1.06	-0.62	-0.19	-0.54	-0.32	1.47	-3.91
	[-0.27,0.05]	[-1.00,3.12]	[-2.77,1.53]	[-0.38,0.00]	[-1.91,0.83]	[-0.60,-0.05]	[0.55,2.40]	[-6.06,-1.75]
FF5	0.04	-0.81	-0.54	0.07	-0.61	-0.06	0.30	-0.13
	[-0.11,0.20]	[-1.91,0.29]	[-1.78,0.71]	[-0.03,0.18]	[-1.37,0.15]	[-0.24,0.11]	[-0.57,1.18]	[-1.17,0.91]
All	-0.03	-0.72	0.54	0.09	-0.16	-0.13	0.92	0.51
	[-0.17,0.12]	[-1.73,0.30]	[-0.63,1.71]	[-0.01,0.19]	[-0.94,0.63]	[-0.31,0.04]	[0.07,1.77]	[-0.42,1.45]
				Panel C: Mimicking				
Mimicking	-0.20	0.02	-0.46	0.01	0.10	-0.07	1.42	0.47
	[-0.27,-0.12]	[-0.46,0.49]	[-0.97,0.06]	[-0.03,0.04]	[-0.13,0.32]	[-0.15,0.02]	[0.83,2.02]	[0.16,0.79]

ison of all possible model combinations. We have 8 traded factors and 8 nontraded factors in total, and we set CAPM as the baseline model, with the market risk factor (MKTRF) always included, as it is difficult to justify that MKTRF would not contribute to pricing with a large cross-section of test assets (although relaxing this constraint is straightforward in our model). This results in $2^{15} = 32,768$ models to evaluate.

Table 6 reports the top 10 marginal likelihood models, the WLS CSR^2, and the H-J distance. It is not surprising that all 8 traded factors are selected because our comparison is not like the GRS test. A traded factor may not have a significant Sharpe ratio increase, but it can still contribute to pricing the cross section of test assets, and most traded factors are not weak (they are themselves excess returns and will correlate more with test assets). The top 1 model has only one additional nontraded factor,

Table 6: **Top 10 Models Ranked by Marginal Likelihood**

This table reports the top 10 models from all possible combinations of 16 factors, out of a total of 32,768 models. The log marginal likelihood, WLS CSR^2 (%), and H-J distance (%) are presented.

Rank	Top 10 Model	logML	CSR^2	H-J D
1	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + HKMcapital	276319	49.10	0.0071
2	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ	276298	49.03	0.0071
3	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PCEND + HKMcapital	276244	50.33	0.0070
4	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PCEND	276221	50.13	0.0070
5	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PCEDG + HKMcapital	276219	49.13	0.0071
6	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + HKMcapital + PEAR	276207	49.66	0.0071
7	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PCEDG	276195	49.10	0.0071
8	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PEAR	276184	49.55	0.0071
9	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + PCEDG + PCEND + HKMcapital	276176	50.54	0.0070
10	MKTRF + SMB + HML + RMW + CMA + UMD + BAB + QMJ + IndProd + HKMcapital	276152	49.21	0.0071

HKMcapital, which is aligned with the findings in the previous section. The CSR^2 is 49.1%, which is not the highest, while the H-J distance is 1.990, which is not the smallest among the top 10 models. Adding more factors will increase CSR^2, as for the top 3 models, although the increase is only about 1%. It will also decrease the H-J distance by only 0.014. The top 2 model is the pure traded factor model, which means that, except for HKMcapital, other nontraded factors are weak or hard to increase pricing performance given a large cross section of test assets.

Table 7 shows the estimation of risk premia in the top 1 models. The rows provide the posterior mean, standard deviation, and credible intervals for the Bayesian approach, along with the mean, standard error, and confidence intervals for WLS and mimicking portfolio estimates. HKMcapital is the only nontraded factor selected in the top 1 model, with the risk premia estimation consistently across different estima-

Table 7: **Risk Premia Estimation in Top 1 Model**

This table presents the risk premia estimates for the top-performing model (see table 6) among all combinations of 16 factors. Estimates are reported using Bayesian, WLS, and mimicking portfolio approaches. For Bayesian estimation, we provide the posterior mean, standard deviation, and the 95% credible interval (lower and upper bounds). For WLS and mimicking portfolios, we report the mean, standard error, and 95% confidence interval. All values are expressed in percentage terms.

	MKTRF	SMB	HML	RMW	CMA	UMD	BAB	QMJ	HKMcapital
					Bayes				
postmean	0.77	0.09	0.11	0.39	0.20	0.38	0.75	0.46	1.44
postsd	0.24	0.14	0.14	0.11	0.10	0.23	0.17	0.12	0.43
lower	0.31	-0.19	-0.17	0.18	0.00	-0.06	0.43	0.22	0.59
upper	1.25	0.37	0.40	0.60	0.41	0.81	1.09	0.71	2.30
					WLS				
mean	0.77	0.10	0.13	0.31	0.15	0.49	0.45	0.24	0.92
std.err	0.21	0.14	0.15	0.12	0.11	0.21	0.22	0.14	0.43
lower	0.36	-0.17	-0.17	0.07	-0.05	0.07	0.02	-0.02	0.07
upper	1.18	0.37	0.43	0.55	0.36	0.90	0.89	0.51	1.77
					Mimicking				
mean	1.05	0.07	0.14	0.21	0.00	0.19	0.02	0.03	1.25
std.err	0.21	0.14	0.15	0.09	0.07	0.10	0.10	0.09	0.26
lower	0.64	-0.20	-0.14	0.03	-0.15	-0.01	-0.18	-0.15	0.73
upper	1.45	0.34	0.43	0.38	0.14	0.38	0.22	0.22	1.76

tion methods and statistically significant.

Figure 1 illustrates the factor selection counts of factors across the top 10, 20, and 30 models, providing insights into their relative importance. Traded factors such as MKT, SMB, and HML exhibit high selection rates, remaining included in nearly all top 30 models. This indicates the robust explanatory power. Among the nontraded factors, HKMcapital demonstrates the highest selection rate, steadily outperforming other nontraded factors. In contrast, factors such as LTY, and M2_SA show the lowest selection rates, indicating limited inclusion and weaker relevance across the model rankings. Factors like PCEDG and PEAR exhibit moderate selection rates. These results confirm that most nontraded factors are weak, with significantly lower explanatory power compared to traded factors in the beta-pricing model.

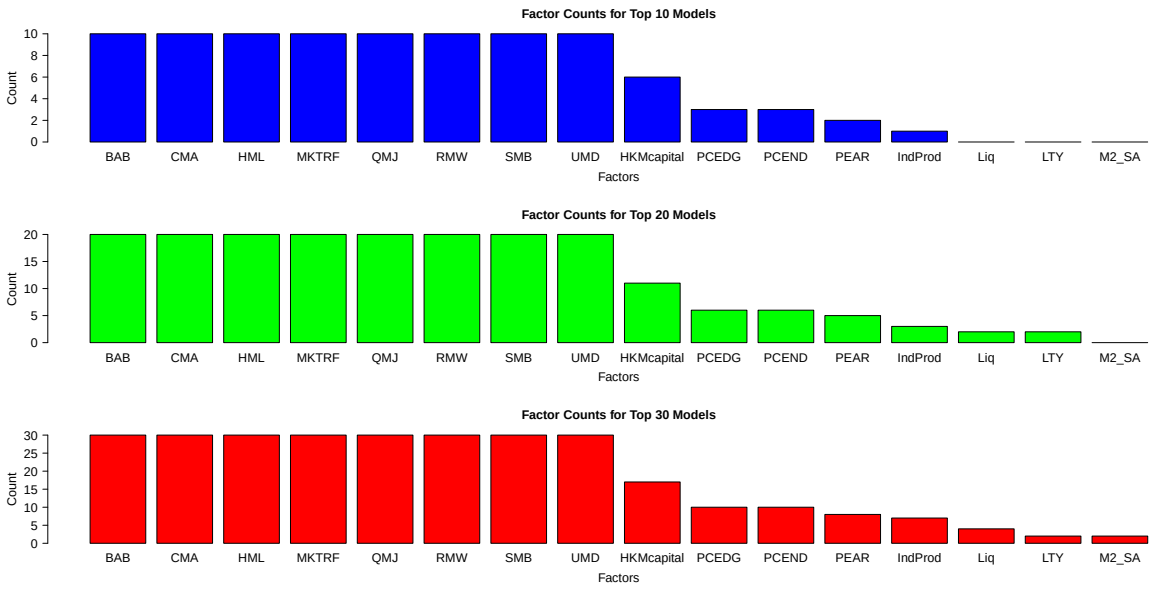


Figure 1: Factor Selection Count

This figure presents a bar plot illustrating the frequency of factor selection among the top 10, 20, and 30 models derived from all possible combinations of 16 factors.

4.5 Factor Model Implied Investment Performance

By selecting the best-performing model from a set of candidate factors, we now investigate how the chosen model facilitates the construction of the tangency portfolio from a set of test assets, particularly when the number of test assets is large. A key advantage of our modeling framework lies in its capacity to estimate the inverse covariance structure of any given number of test assets by leveraging the covariance structure of the factors. This feature enables us to construct the tangency portfolio using the estimated model-implied means and covariances of the test assets. Crucially, the model selected through our Bayesian procedure is expected to best explain the cross-section of test assets. Consequently, it is anticipated to provide more accurate model-implied estimates of both the mean and covariance, thereby improving the effectiveness of the resulting portfolio construction.

We divide the entire sample period into an in-sample (INS) period (the first 200 months) and an out-of-sample (OOS) period (the remaining months). Model selection is performed on the INS data to identify the top-performing model. Then for each model estimation, including top models and benchmark FF3 and FF5 or even all fac-

tors, the expected return implied by the selected model is given by $\boldsymbol{\mu}_{tan} = \boldsymbol{B}[\boldsymbol{\mu}^T, \boldsymbol{\lambda}^{NT}]$ ($\boldsymbol{\mu}_{tan} = \boldsymbol{B}\boldsymbol{\lambda}$ for WLS), while the inverse of the covariance matrix is computed as:

$$\Sigma_{tan}^{-1} = \Sigma_e^{-1} - \Sigma_e^{-1}\boldsymbol{B}(\Sigma_u^{-1} + \boldsymbol{B}'\Sigma_e^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}'\Sigma_e^{-1}$$

where the Σ_e^{-1} is a diagonal matrix. In this case, we can avoid inverting a high-dimensional matrix by utilizing the factor covariance structure. The weight of the tangency portfolio is given by $\frac{\Sigma_{tan}^{-1}\boldsymbol{\mu}_{tan}}{\mathbf{1}^T\Sigma_{tan}^{-1}\boldsymbol{\mu}_{tan}}$.

Table 8 reports the performance of models based on the Sharpe Ratios of tangency portfolios constructed from the test assets for both Bayes and WLS. The top1, top2, and top3 models, described in Panel C, are selected following a procedure analogous to that in Section 4.4, but relying exclusively on the INS period. Each row in the panel corresponds to a specific factor model used to estimate portfolio weights, which are subsequently applied to construct tangency portfolios. All estimations are conducted solely using INS data. Model performance during the INS period reflects tangency portfolios constructed with INS returns, whereas performance during the OOS period is evaluated by applying the tangency portfolio weights estimated from the INS period to OOS data. We report results for test asset returns across two datasets: the full set of 302 test assets in Panel A and the FF25 portfolios in Panel B.

Considering the Bayes INS results, we observe that the top 1 model provides better tangency portfolio weights for both all 302 test assets and FF25, achieving annualized Sharpe Ratios of 2.55 and 1.97, respectively. It also significantly outperforms the standard FF3 and FF5 models. Notably, increasing the number of factors, as shown in the row "ALL Factors," does not improve performance, even in sample. When turning to the Bayes OOS results, the performance remains highly consistent. The top 1 model continues to outperform others, achieving an OOS tangency portfolio Sharpe Ratio of 0.87 for all test assets and 0.72 for FF25.

For all portfolios, the WLS results differ as they favor all factors with the highest INS and OOS Sharpe Ratios. However, these Sharpe Ratios remain lower than the Bayes Sharpe Ratio of the top1 model. This highlights the importance of our Bayesian approach, not only for model comparison but also for model estimation. For FF25, the

WLS OOS Sharpe Ratio of top1 model is slightly higher than that of Bayes, indicating that WLS performs well in small cross-sections. Nonetheless, its performance declines as the number of test assets increases, suggesting that the superior performance of our Bayesian estimation approach is most pronounced in high-dimension.

In summary, the INS and OOS results further support the conclusion that models selected through our Bayesian ML approach tend to exhibit stronger explanatory power for test asset returns economically, which is reflected in the model-implied mean and covariance structure estimated for tangency portfolio construction.

Table 8: **In Sample and Out-of-Sample Model Performance**

This table reports the annualized Sharpe Ratios of the tangency portfolios, estimated using six factor models, for both in-sample (INS) and out-of-sample (OOS) test assets. The estimation methods include Bayes and WLS. The test assets include all test portfolios (Panel A) as well as the Fama-French 25 portfolios (Panel B). Panel C presents the top three models selected based on their in-sample ML performance, evaluated using data from the in-sample period of all portfolios.

	Bayes INS SR	WLS INS SR	Bayes OOS SR	WLS OOS SR
	Panel A: All Port			
top1	2.545	1.413	0.866	0.667
top2	2.403	1.480	0.821	0.624
top3	2.020	1.395	0.737	0.645
all	1.945	1.554	0.794	0.782
FF3	0.901	0.918	0.706	0.662
FF5	1.316	1.280	0.839	0.759
	Panel B: 25 Size-Value			
top1	1.965	1.626	0.724	0.735
top2	1.831	1.667	0.716	0.718
top3	1.576	1.663	0.516	0.670
all	1.608	1.715	0.652	0.735
FF3	0.816	0.888	0.457	0.460
FF5	1.096	1.211	0.693	0.722
	Panel C: Top Models			
top1	FF3+RMW+CMA+UMD+QMJ+PCEDG+HKMcapital+PEAR			
top2	FF3+RMW+UMD+BAB+QMJ+PCEDG+PCEND+PEAR			
top3	FF3+RMW+UMD+BAB+PCEDG			

5 Revisiting Kan, Robotti, and Shanken (2013)

Kan et al. (2013) evaluate the asset pricing performance through CSR^2 and derive the asymptotic distribution of this measure. They show that many large CSR^2 differences are not statistically significant. Based on our approach, we can revisit their results and confirm this insignificant since the insignificant increase of CSR^2 naturally translates to our marginal likelihood decrease when adding new useless factors.

They consider 8 asset pricing models with a sample period ranging from February 1959 to July 2007. The models and the cross-sectional specifications are as follows,

1. CAPM

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{vw}\boldsymbol{\lambda}_{vw}$$

2. The conditional CAPM (C-LAB) of Jagannathan and Wang (1996)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{vw}\boldsymbol{\lambda}_{vw} + \boldsymbol{\beta}_{lab}\boldsymbol{\lambda}_{lab} + \boldsymbol{\beta}_{\text{prem}}\boldsymbol{\lambda}_{\text{prem}}$$

3. FF3

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{vw}\boldsymbol{\lambda}_{vw} + \boldsymbol{\beta}_{smb}\boldsymbol{\lambda}_{smb} + \boldsymbol{\beta}_{hml}\boldsymbol{\lambda}_{hml}$$

4. ICAPM (Petkova, 2006)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{vw}\boldsymbol{\lambda}_{vw} + \boldsymbol{\beta}_{\text{term}}\boldsymbol{\lambda}_{\text{term}} + \boldsymbol{\beta}_{\text{def}}\boldsymbol{\lambda}_{\text{def}} + \boldsymbol{\beta}_{\text{div}}\boldsymbol{\lambda}_{\text{div}} + \boldsymbol{\beta}_{\text{rf}}\boldsymbol{\lambda}_{\text{rf}}$$

5. The unconditional consumption model (CCAPM)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{cg}\boldsymbol{\lambda}_{cg}$$

6. CC-CAY from Lettau and Ludvigson (2001)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{cay}\boldsymbol{\lambda}_{cay} + \boldsymbol{\beta}_{cg}\boldsymbol{\lambda}_{cg} + \boldsymbol{\beta}_{cg\cdot cay}\boldsymbol{\lambda}_{cg\cdot cay}$$

7. U-CCAPM from Parker and Julliard (2005)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{cg36}\boldsymbol{\lambda}_{cg36}$$

8. D-CCAPM from Yogo (2006)

$$E(\boldsymbol{R}) = \boldsymbol{\beta}_{vw}\boldsymbol{\lambda}_{vw} + \boldsymbol{\beta}_{cg}\boldsymbol{\lambda}_{cg} + \boldsymbol{\beta}_{cgdur}\boldsymbol{\lambda}_{cgdur}$$

In summary, there are 14 factors, comprising 2 traded factor models, 3 nontraded factor models, and 3 mixed models. Details of the data are provided in Appendix A.I.1. Table 9 presents the estimation results, log marginal likelihoods, and sample CSR^2 for these 8 models. When the zero-beta rate is constrained to the risk-free rate, we focus on excess return analysis as in Kan et al. (2013). The CSR^2 reported is the uncentered measure used in Kan et al. (2013) under a specified zero-beta rate.[4] GLS CSR^2 uses a weighting matrix of the inverse of the covariance matrix of return (as in Kan et al. (2013)), while GLS CSR^2_{res} uses the inverse of the residual covariance matrix.

Our findings on OLS and GLS estimation and CSR^2 align closely with Kan et al. (2013). However, for their best-performing model, ICAPM, slight deviations in estimation and CSR^2 emerge, likely due to updated data or differing processing methods. These discrepancies are minor.[5] In Kan et al. (2013), although ICAPM shows slightly higher CSR^2 than FF3, this CSR^2 difference is not statistically significant.

FF3 shows the best performance regarding marginal likelihood, much higher than others. The CSR^2 is the highest in our sample, while only slightly higher than that of ICAPM. However, the marginal likelihood for ICAPM is much lower, falling below that of the single-factor CAPM. Beyond the traded-only model FF3, D-CCAPM, which considers both nondurable and durable consumptions, achieves the second-best model. This is aligned with our main results, which show that two consumption measures are the second best-nontraded factors after HKMcapital. In Section 6, we revisit 5 popular consumption measures in detail.

In addition to 8 pre-specified models, we then consider all possible model combinations of factors within these 8 models. Table 10 shows the top 5 models of all combinations from Kan et al. (2013). It includes the risk premia estimation under Bayes, OLS,

[4]The use of uncentered CSR^2 doesn't change the relative ranking, but only to make sure the CSR^2 always larger than 0.

[5]Even for the standard FF3 model, the results differ slightly due to data updates, but this does not affect the overall conclusion.

Table 9: **Comparison of 8 Models in Kan, Robotti, and Shanken (2013)**

This table summarizes the eight models from Kan et al. (2013). The models are estimated using monthly returns of the 25 Fama–French size and book-to-market portfolios and five industry portfolios, covering February 1959 to July 2007. For Bayesian estimation, we present the posterior mean, standard deviation, and 95% credible interval bounds. For OLS and GLS, we report the mean and standard error. All figures are in percentages. Each model's log marginal likelihood and uncentered CSR^2 are reported. GLS CSR^2 employs the return covariance matrix as a weight (Kan et al., 2013), while CSR^2_{res} uses the residual covariance matrix.

		Bayes				OLS		GLS	
		postmean	postsd	lower	upper	mean	se	mean	se
Top 1 : FF3	vw	0.32	0.22	-0.12	0.74	0.49**	0.18	0.51**	0.18
	smb	-0.08	0.16	-0.38	0.24	-0.08	0.13	-0.08	0.13
	hml	0.37	0.13	0.10	0.62	0.39**	0.12	0.44**	0.11
		logML				CSR^2		CSR^2	CSR^2_{res}
		58419				95.83		29.69	97.34
Top 2 : D_CCAPM	vw	0.30	0.22	-0.13	0.71	0.54**	0.18	0.51**	0.18
	cg	1.13	0.20	0.77	1.57	1.05**	0.29	1.05**	0.15
	cgdur	0.29	0.42	-0.53	1.13	0.29	0.61	0.29	0.40
		logML				CSR^2		CSR^2	CSR^2_{res}
		51904				89.27		7.51	91.89
Top 3 : CAPM	vw	0.30	0.22	-0.12	0.73	0.63**	0.19	0.51**	0.18
		logML				CSR^2		CSR^2	CSR^2_{res}
		51832				86.16		6.04	91.75
Top 4 : ICAPM	vw	0.32	0.21	-0.11	0.72	0.58**	0.18	0.54**	0.18
	term	0.23	0.07	0.09	0.37	0.39**	0.08	0.18**	0.04
	def	0.02	0.04	-0.05	0.10	0.02	0.05	0.02	0.03
	div	0.03	0.01	0.01	0.06	0.03	0.01	0.03	0.01
	rf	-0.59	0.15	-0.89	-0.30	-0.70**	0.14	-0.30**	0.09
		logML				CSR^2		CSR^2	CSR^2_{res}
		51768				94.88		22.41	93.26
Top 5 : C_LAB	vw	0.30	0.22	-0.13	0.71	0.55**	0.18	0.50**	0.18
	lab	-0.48	0.08	-0.65	-0.34	-0.48	0.13	-0.12**	0.06
	prem	-0.02	0.09	-0.18	0.16	0.45**	0.13	0.45**	0.07
		logML				CSR^2		CSR^2	CSR^2_{res}
		51755				89.24		9.00	92.07
Top 6 : CCAPM	cg	0.64	0.07	0.51	0.77	0.66**	0.20	0.31**	0.12
		logML				CSR^2		CSR^2	CSR^2_{res}
		42715				88.76		4.63	4.96
Top 7 : U_CCAPM	cg36	3.81	0.42	2.99	4.64	4.93**	1.36	1.83**	0.52
		logML				CSR^2		CSR^2	CSR^2_{res}
		42625				95.04		9.42	10.17
Top 8 : CC_CAY	cg	0.62	0.07	0.47	0.76	0.64**	0.19	0.48**	0.13
	cay	0.20	0.61	-1.03	1.43	0.20	0.83	1.39**	0.35
	cgcay	0.00	0.01	-0.01	0.01	0.00	0.01		0.00
		logML				CSR^2		CSR^2	CSR^2_{res}
		42612				89.51		18.29	18.94

and GLS, and log marginal likelihood and CSR^2. Since all models contain FF3 with additional nontraded factors, we only report risk premia estimates for nontraded factors

Table 10: **Comparison of All Model Specification**

This table presents the top 5 models among all combinations of 14 factors analyzed in Kan et al. (2013). For Bayesian models, we report the posterior mean, standard deviation, and the 95% credible interval bounds. For OLS and GLS, we provide the mean and standard error. All figures are in percentages. Each model includes FF3, so we report only the risk premia estimates for the nontraded factor. Additionally, we present the log marginal likelihood and uncentered CSR^2. GLS CSR^2 uses the return covariance matrix as a weighting matrix (Kan et al., 2013), while CSR^2_{res} employs the residual covariance matrix.

		Bayes				OLS		GLS	
		postmean	postsd	lower	upper	mean	se	mean	se
Top 1	cg	0.26	0.90	-1.29	1.33	-0.34	0.21	-0.10	0.16
		logML				CSR^2		CSR^2	CSR^2_{res}
		58459				95.96		30.16	97.35
Top 2		logML				CSR^2		CSR^2	CSR^2_{res}
		58419				95.83		29.69	97.34
Top 3	cgdur	3.02	0.69	1.78	4.46	1.60**	0.67	-0.09	0.51
		logML				CSR^2		CSR^2	CSR^2_{res}
		58399				96.1		29.73	97.34
Top 4	cgcay	-0.01	0.01	-0.03	0.02	-0.01	0.00	0.00	0.00
		logML				CSR^2		CSR^2	CSR^2_{res}
		58396				96.48		30.41	97.37
Top 5	def	-0.06	0.06	-0.17	0.05	0.07	0.05	0.00	0.04
		logML				CSR^2		CSR^2	CSR^2_{res}
		58394				95.99		29.69	97.34

in different models. Although the FF3 model shows the best performance among the 8 models, it can be enhanced by adding another nontraded factor. The best model adds cg into the FF3, with a log marginal likelihood achieving 58459, although its risk premia is not significant. However, FF3 remains the second-best model, with all other models incorporating nontraded factors ranking below it. This supports our main empirical analysis, indicating that most of these nontraded factors are weak and fail to provide any improvement over the FF3 model.

6 Revisiting Consumption-based Model

In Sections 4 and 5, we evaluate factor models tied to consumption, often based on NIPA-reported measures. However, consumption measures are weakly correlated

with test asset returns. To address this, the literature has proposed alternative measures. The validity of these measures in asset pricing is typically assessed via two methods: (i) the two-pass regression approach for linear models, where expected returns are derived from the beta-pricing model, and (ii) the generalized method of moments (GMM) for nonlinear models, leveraging the SDF representation (Hansen, 1982). Kleibergen and Zhan (2020) introduce robust tests for identifying risk premia across different consumption measures. We contribute a Bayesian perspective, evaluating the weakness of these measures through marginal likelihood.

We analyze five consumption measures as in Kleibergen and Zhan (2020): (i) NIPA's reported consumption expenditure on nondurable goods and services ("Reported"), (ii) the three-year consumption measure of Parker and Julliard (2005) ("P-J"), (iii) the fourth-quarter to fourth-quarter consumption measure of Jagannathan and Wang (2007) ("Q4-Q4"), (iv) the garbage-based measure of Savov (2011) ("Garbage"), and (v) the unfiltered NIPA consumption measure of Kroencke (2017) ("Unfiltered"). Further data details are in Appendix A.I.2. The initial 12 years of data are used as training samples.

6.1 Comparison of Consumption Models with MKTRF as test assets

We analyze the model specification in Kleibergen and Zhan (2020)'s original framework, where MKTRF serves as the test asset rather than a factor. The test assets include 30 portfolios sorted by size, value, and investment, resulting in 31 test assets alongside 5 consumption-based nontraded factors. Model details are provided in Appendix A.IV. Table 11 reports estimation results and marginal likelihoods for the 5 consumption models using MKTRF as the test asset.

Among the five measures, the "Garbage" measure outperforms, achieving the highest log marginal likelihood (logML = 1027) and delivering consistent risk premia estimates across Bayesian, OLS, and GLS methods. The Bayesian approach yields a posterior mean risk premium of 2.08 with a tight 95% credible interval (1.26, 3.03), underscoring statistical significance. While other models vary in performance, none surpass the "Garbage" measure in logML.

Table 11: **Estimation and Marginal Likelihood with MKTRF as test assets**

This table summarizes the five consumption models from Kleibergen and Zhan (2020). The test assets include 30 portfolios sorted by size, value, and investment, along with the market portfolio, using annualized data from 1960 to 2014. MKTRF is treated as a test asset rather than a factor. For the Bayesian estimates, we report the posterior mean, standard deviation, and 95% credible interval. For OLS and GLS, we present the mean and standard error. We focus on the risk premia estimate (%) of the nontraded factor. Each model's log marginal likelihood and uncentered CSR^2 are reported. GLS CSR^2 employs the return covariance matrix as weights, while CSR^2_{res} uses the residual covariance matrix.

		Bayes				OLS		GLS	
		postmean	postsd	lower	upper	mean	se	mean	se
Top 1	Garbage	2.08	0.46	1.26	3.03	1.92**	0.76	1.12**	0.43
		logML				CSR^2		CSR^2	CSR^2_{res}
		1027				94.16		8.47	40.28
Top 2	Q4-Q4	1.98	0.31	1.38	2.61	2.39**	0.90	0.55**	0.26
		logML				CSR^2		CSR^2	CSR^2_{res}
		870				96.3		5.68	13.23
Top 3	Unfiltered	4.03	0.66	2.81	5.36	5.27**	1.94	1.18**	0.47
		logML				CSR^2		CSR^2	CSR^2_{res}
		852				86.11		7.57	15.02
Top 4	P-J	4.56	0.66	3.29	5.86	10.00**	3.52	1.99**	0.55
		logML				CSR^2		CSR^2	CSR^2_{res}
		838				80.63		16.24	31.91
Top 5	Reported	-2.66	0.49	-3.65	-1.78	-3.35**	1.55	0.15	0.22
		logML				CSR^2		CSR^2	CSR^2_{res}
		823				39.89		0.55	1.48

These results suggest that under the original setup with only consumption into the model, the "Garbage" is the best one under our evaluation. This aligns with Kleibergen and Zhan (2020), where they state that "except for the 'Garbage' consumption measure, we cannot trust the FM t-test for the risk premia on any of the consumption measures." However, without a traded factor benchmark, we can not evaluate the weakness of consumption measures. Thus, in the next Section, we consider the multifactor setup in Kleibergen and Zhan (2020) to replace MKTRF as a traded factor.

Table 12: **Estimation and Marginal Likelihood under MKTRF Control**

This table summarizes five consumption models from Kleibergen and Zhan (2020). The test assets are 30 portfolios sorted by size, value, and investment, using annualized data from 1960 to 2014. MKTRF is treated as a traded factor. For the Bayesian estimates, we report the posterior mean, standard deviation, and 95% credible interval. For OLS and GLS, we present the mean and standard error. All figures are in percentages. Each model includes the log marginal likelihood and uncentered CSR^2. GLS CSR^2 uses the return covariance matrix as weights, while CSR^2_{res} employs the residual covariance matrix.

		Bayes				OLS		GLS	
		postmean	postsd	lower	upper	mean	se	mean	se
Top 1	MKTRF	5.00	2.66	-0.12	10.11	5.62**	2.37	4.30	2.29
		logML				CSR^2		CSR^2	CSR^2_{res}
		1872				89.92		4.32	84.73
Top 2	MKTRF	5.07	2.59	-0.05	10.02	5.96**	2.39	4.31	2.29
	Unfiltered	1.90	0.59	0.79	3.12	1.44	0.83	0.62	0.49
		logML				CSR^2		CSR^2	CSR^2_{res}
		1853				91.56		5.63	84.95
Top 3	MKTRF	5.09	2.56	-0.10	9.84	5.62**	2.39	4.30	2.29
	Garbage	2.34	0.67	1.12	3.78	0.56	0.99	0.74	0.46
		logML				CSR^2		CSR^2	CSR^2_{res}
		1843				89.92		5.08	84.86
Top 4	MKTRF	5.09	2.58	-0.03	10.01	5.85**	2.40	4.30	2.29
	Q4-Q4	1.06	0.32	0.47	1.70	0.56	0.45	0.18	0.25
		logML				CSR^2		CSR^2	CSR^2_{res}
		1842				90.51		4.45	84.89
Top 5	MKTRF	5.09	2.58	-0.04	9.97	7.07**	2.48	4.33	2.29
	P-J	2.05	0.58	0.98	3.17	3.57**	1.25	0.74	0.54
		logML				CSR^2		CSR^2	CSR^2_{res}
		1815				93.8		5.92	85.04
Top 6	MKTRF	5.08	2.58	-0.06	10.03	5.68**	2.40	4.31	2.29
	Reported	0.68	0.28	0.15	1.26	0.07	0.40	-0.17	0.22
		logML				CSR^2		CSR^2	CSR^2_{res}
		1805				89.94		5.02	84.86

6.2 Comparison of Consumption Models under MKTRF control

Another model specification of the consumption-based model is to treat MKTRF as the only traded factor control. The test assets are 30 test portfolios sorted by size, value, and investment. Table 12 reports the single factor comparison against MKT.

The results indicate that, under this specification, all measures appear relatively

weak, suggesting that the consumption growth alone struggles to explain the variation in returns without additional support from other traded factors. Despite this, the MKT-only model achieves the highest marginal likelihood. This conclusion is consistent with Kleibergen and Zhan (2020), which uses a F test of identification to show that all consumption measures are weak. They can not reject that all five consumption measures' $\boldsymbol{\beta} = 0$.

Within our Bayesian method, we reexamine the consumption measure and achieve the same conclusion that if the data is noisy, such as consumption, the meaningful inference is infeasible. Given the limited information in the data, we cannot determine whether the consumption growth factor accounts for part of the variation in the cross section of expected asset returns, nor can we rule out the possibility that it explains all of the variations.

7 Conclusion

Beta-pricing models are widely used for estimating and testing asset pricing models, yet key empirical challenges persist. First, standard estimation methods primarily focus on evaluating a risk factor's relevance, while Bayesian approaches better address parameter uncertainty. Second, comparisons of beta-pricing models often rely on CSR^2 and the Hansen-Jagannathan distance, both of which have ambiguous statistical significance. Third, while economic constraints are theoretically essential, they are frequently neglected in the literature. A Bayesian framework, as proposed in Pastor and Stambaugh (2003), offers a robust approach to addressing these issues.

We develop a Bayesian model that enables the estimation and comparison of both traded and non-traded factors, while closely adhering to the guidelines established by Lewellen et al. (2010) to impose several asset pricing constraints. The model's performance is evaluated based on cross-sectional pricing ability and assessed via marginal likelihood. Simulation results indicate that weak factors are excluded even when their associated risk premia are significant. Marginal likelihood effectively selects the optimal model among the factors, whereas CSR^2 and the Hansen-Jagannathan distance

are more challenging to differentiate.

Through empirical examination of 16 traded and non-traded factors, we find that traded factors significantly explain a substantial cross-sectional return variation, whereas most non-traded factors exhibit weakness. The intermediary capital ratio (He et al., 2017) is the only non-traded factor in the top model, which exhibits significant risk premia. Evaluting through tangency portfolio constructed from test assets, we find that the in-sample top model outperforms both in-sample and out-of-sample. Beyond model selection, Bayesian estimation also demonstrates superior performance, particularly in large cross-sections. In revisiting Kan et al. (2013), we find the Fama-French three-factor model outperforms eight pre-specified models, with nondurable consumption growth further improving its fit. Examining consumption-based models from Kleibergen and Zhan (2020), we observe weak performance across all consumption measures, with garbage consumption growth exhibiting the strongest results.

Appendix

A.I Data in Kan et al. (2013) and Kleibergen and Zhan (2020)

A.I.1 Data in Kan et al. (2013)

There are 14 factors in Kan et al. (2013), including

1. MKTRF, SMB, HML, directly downloaded from French's website.

2. prem, the lagged yield spread between Baa- and Aaa-rated corporate bonds. We use data from Amit Goyal's website, and the original source is from the Board of Governors of the Federal Reserve System.

3. lab, the growth rate in per capita labor income, L, defined as the difference between total personal income and dividend payments, divided by the total population (from FRED). Following Kan et al. (2013), we use a 2-month moving average to construct the growth rate lab = $(L_{t-1} + L_{t-2})/(L_{t-2} + L_{t-3}) - 1$, for the purpose of minimizing the influence of measurement error.

4. term, the difference between the yields of 10- and 1-year government bonds, which are from the Board of Governors of the Federal Reserve System. def, the difference between the yields of long-term corporate Baa bonds and long-term government bonds, which are still from Amit Goyal's website. div, the dividend yield, using dividend from Amit Goyal's website divided by lagged S& P 500 Index. [6] rf is the 1-month T-bill yield (from CRSP, Fama Risk Free Rates). The actual factors of term, def, div, and rf are their innovations from a VAR(1) system of seven state variables that also include MKTRF, SMB, and HML.

5. cg, the growth rate in real per capita nondurable consumption (seasonally adjusted at annual rates). It is calculated using personal consumption expenditures, nondurable goods (chain-type price index) and population from FRED.

[6] Kan et al. (2013)use Center for Research in Security Prices (CRSP) value weighted stock market portfolio for dividend, which may be the cause of the slight difference.

6. cay, the conditioning variable, is a consumption-aggregate wealth ratio from Lettau and Ludvigson (2001). We linearly interpolate the quarterly values of cay to permit analysis at the monthly frequency. cgcay is just the product of cg and cay.

7. cg36, the growth rate in real per capita nondurable consumption over 3 years starting with the given month.

8. cgdur, the growth rate in real per capita durable consumption (seasonally adjusted at annual rates). It is calculated similarly as cg.

A.I.2 Data in Kleibergen and Zhan (2020)

5 consumption measures are used in Kleibergen and Zhan (2020). Thanks for Kroencke (2017) making four of them publicly available. Garbage measures are downloaded from the U.S. Environmental Protection Agency (EPA) 's website. According to Savov (2011), yard trimmings are removed because they appear to be poorly measured. Regarding test assets, rather than commonly used Fama-French portfolios, they use univariate sorted to size, value, and investment portfolios, which can also be downloaded from Kroencke's website.

A.II Additional Tables and Figures

Table A.1: **Definition of Factors**

This table shows the factor explanation and sources. We include 16 factors in total, with 8 traded factors and 8 nontraded factors.

Factors	Explanations	Sources
	Traded Factors	
MKTRF	The excess return of the market portfolio (the value-weighted return of all CRSP firms minus the one-month Treasury bill rate)	Sharpe (1964)
SMB	Small Minus Big (the spread of average returns on nine small and nine big stock portfolios)	Fama and French (2015)
HML	High Minus Low (the spread of average returns on two value and two growth portfolios)	Fama and French (2015)
RMW	Robust Minus Weak (the spread of average returns on two robust and two weak operating profitability portfolios)	Fama and French (2015)
CMA	Conservative Minus Aggressive (the spread of average returns on two conservative and two aggressive investment portfolios)	Fama and French (2015)
MOM	The momentum factor (the average of a small stock spread factor and a big stock spread factor)	Fama and French (2018)
BAB	Betting Against Beta (long leveraged low-beta assets and short high-beta assets)	Frazzini and Pedersen (2014)
QMJ	Quality Minus Junk (long high-quality assets and short low-quality assets)	Asness et al. (2019)
	Nontraded Factors	
IndProd	Industrial Production	Chen et al. (1986)
Liq	Liquidity	Pastor and Stambaugh (2003)
LTY	Long Term Yield	Sweeney and Warga (1986)
M2_SA	M2 Seasonally Adjusted	Chan et al. (1996)
PCEDG	Personal Consumption Expenditures Durable	Yogo (2006)
PCEND	Personal Consumption Expenditures Non-Durable	Yogo (2006)
HKMcapital	Intermediary Capital Ratio and Risk Factor	He et al. (2017)
PEAR	Monthly presidential economic approval rating index	Chen et al. (2023)

Table A.2: **Rejection Rate in Simulation**

This table shows the 5% rejection rate of the frequency of rejecting the null hypothesis $H_0 : \lambda = \lambda_{true}$ in simulation under the true model. There are 8 factors, including 6 traded factors and 2 nontraded factors. The parameters are calibrated from the true test assets and true factors. The number is calculated from 1000 simulation runs.

	WLS	Bayes	BayesNC	WLS	Bayes	BayesNC
	25 test assets			100 test assets		
MKTRF	0.055	0.046	0.076	0.050	0.041	0.353
SMB	0.052	0.060	0.066	0.056	0.043	0.103
HML	0.041	0.040	0.045	0.044	0.039	0.064
RMW	0.047	0.057	0.036	0.055	0.049	0.069
CMA	0.060	0.038	0.039	0.037	0.032	0.042
UMD	0.069	0.075	0.047	0.054	0.064	0.065
PCEDG	0.049	0.042	0.020	0.060	0.039	0.039
HKMcapital	0.069	0.041	0.042	0.050	0.043	0.152

Table A.3 **Estimation for Traded Factors under Different Test Assets**

This table reports estimates from the traded-only model using various test assets. We compare Bayesian models alongside WLS with and without traded factor constraints. The 95% credible (confidence) intervals are shown in parentheses.

	MKTRF	SMB	HML	RMW	CMA	UMD	BAB	QMJ
					5×5 Size and Book-to-Market			
Bayes No Constraint	0.88	0.13	0.07	0.38	0.44	2.4	0.38	0.25
	[0.37,1.34]	[-0.19,0.45]	[-0.27,0.39]	[-0.01,0.77]	[-0.28,1.18]	[0.23,4.62]	[-1.16,1.93]	[-0.31,0.84]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.81	0.10	0.12	0.40	0.28	2.26	0.29	0.41
	[0.40,1.22]	[-0.17,0.37]	[-0.17,0.41]	[0.01,0.80]	[-0.27,0.82]	[0.80,3.72]	[-0.66,1.24]	[-0.02,0.84]
					17 Industry			
Bayes No Constraint	0.72	-0.5	-0.02	0.3	0.03	0.39	-0.32	0.07
	[0.26,1.17]	[-1.56,0.51]	[-0.47,0.46]	[-0.24,0.88]	[-0.45,0.52]	[-1.43,2.21]	[-1.86,1.17]	[-0.46,0.60]
Bayes With Constraint	0.71	0.08	0.12	0.4	0.21	0.43	0.74	0.48
	[0.26,1.18]	[-0.21,0.35]	[-0.16,0.40]	[0.19,0.60]	[0.02,0.41]	[0.01,0.85]	[0.39,1.07]	[0.25,0.71]
WLS	0.80	-0.44	-0.09	0.30	0.00	0.54	-0.05	0.11
	[0.38,1.21]	[-1.11,0.23]	[-0.48,0.31]	[-0.14,0.73]	[-0.47,0.47]	[-0.65,1.73]	[-1.06,0.95]	[-0.34,0.56]
					5×5 Operating Profitability and Investment			
Bayes No Constraint	0.81	0.08	0.24	0.42	0.18	0.7	-0.21	0.01
	[0.36,1.25]	[-0.76,0.93]	[-0.26,0.74]	[0.03,0.80]	[-0.07,0.46]	[-0.49,1.93]	[-1.49,0.95]	[-0.48,0.48]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.77	0.03	0.17	0.37	0.23	0.41	-0.19	0.06
	[0.36,1.18]	[-0.61,0.67]	[-0.27,0.61]	[0.05,0.70]	[0.01,0.44]	[-0.54,1.35]	[-1.09,0.71]	[-0.33,0.46]
					5×5 Size and Variance			
Bayes No Constraint	0.84	0.2	-0.36	0.07	-0.51	0.39	1.37	0.27

Table A.3 **continued from previous page**

	MKTRF	SMB	HML	RMW	CMA	UMD	BAB	QMJ
	[0.36,1.30]	[-0.18,0.60]	[-1.68,0.99]	[-0.38,0.56]	[-1.33,0.29]	[-1.69,2.43]	[0.17,2.60]	[-0.41,0.96]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.80	0.14	-0.75	0.26	-0.95	-0.16	1.30	0.62
	[0.39, 1.22]	[-0.18, 0.46]	[-1.65, 0.15]	[-0.12, 0.65]	[-1.46,-0.43]	[-1.58, 1.27]	[0.44, 2.16]	[0.14, 1.09]
	5×5 Size and Net Share Issues							
Bayes No Constraint	0.68	0	0.62	0.13	0.72	3.48	0.77	0.08
	[0.32,1.03]	[-0.32,0.31]	[-0.08,1.38]	[-0.35,0.64]	[0.29,1.16]	[1.94,5.25]	[-0.37,1.91]	[-0.38,0.54]
Bayes With Constraint	0.71	0.08	0.12	0.4	0.21	0.42	0.74	0.47
	[0.28,1.17]	[-0.21,0.35]	[-0.17,0.41]	[0.19,0.61]	[0.01,0.42]	[0.00,0.85]	[0.39,1.07]	[0.25,0.71]
WLS	0.79	0.03	0.47	0.20	0.44	2.65	0.78	0.21
	[0.38,1.21]	[-0.26,0.31]	[-0.04,0.98]	[-0.18,0.59]	[0.08,0.79]	[1.37,3.93]	[0.08,1.49]	[-0.14,0.56]
	5×5 Size and Accruals							
Bayes No Constraint	0.82	0.06	-0.3	-0.28	0.13	1.07	-0.45	0.11
	[0.34,1.30]	[-0.27,0.41]	[-1.08,0.47]	[-0.85,0.30]	[-0.38,0.66]	[-1.14,3.19]	[-2.05,1.11]	[-0.39,0.64]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.80	0.00	0.01	-0.13	0.29	1.05	-0.44	0.14
	[0.38,1.23]	[-0.29,0.29]	[-0.69,0.72]	[-0.60,0.35]	[-0.08,0.65]	[-0.45,2.56]	[-1.56,0.67]	[-0.28,0.56]
	5×5 Size and Beta							
Bayes No Constraint	0.86	0.17	-0.08	0.1	-0.26	1.79	0.88	-0.33
	[0.38,1.34]	[-0.18,0.51]	[-1.10,0.92]	[-0.32,0.51]	[-0.88,0.35]	[0.11,3.52]	[0.15,1.62]	[-0.87,0.19]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.81	0.13	-0.12	0.15	-0.36	1.69	0.77	-0.10
	[0.39,1.23]	[-0.15,0.42]	[-0.72,0.48]	[-0.21,0.52]	[-0.78,0.07]	[0.27,3.12]	[0.19,1.34]	[-0.57,0.36]

Table A.3 **continued from previous page**

	MKTRF	SMB	HML	RMW	CMA	UMD	BAB	QMJ
				5 × 5 Size and Momentum				
Bayes No Constraint	0.9	0.16	0.23	-0.09	0.75	0.43	0.32	0.24
	[0.41,1.37]	[-0.19,0.51]	[-1.06,1.68]	[-0.65,0.46]	[-0.17,1.72]	[-0.06,0.90]	[-1.03,1.57]	[-0.32,0.82]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.85	0.07	0.67	-0.21	0.80	0.53	-0.19	0.19
	[0.43,1.27]	[-0.22,0.36]	[-0.22,1.56]	[-0.69,0.26]	[0.11,1.48]	[0.12,0.95]	[-1.15,0.78]	[-0.26,0.63]
				5 × 5 Size and Operating Profitability				
Bayes No Constraint	0.82	0.12	0.34	0.34	0.44	0.97	-0.39	-0.09
	[0.31,1.29]	[-0.21,0.44]	[-0.17,0.91]	[0.10,0.60]	[-0.12,1.06]	[-0.87,2.90]	[-1.90,0.97]	[-0.58,0.39]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.77	0.03	0.38	0.35	0.23	1.05	-0.17	-0.03
	[0.36,1.18]	[-0.24,0.31]	[-0.08,0.83]	[0.12,0.58]	[-0.21,0.67]	[-0.22,2.31]	[-1.09,0.76]	[-0.43,0.38]
				5 × 5 Size and Investment				
Bayes No Constraint	0.86	-0.11	1.31	-0.64	0.08	1.55	0.24	-0.22
	[0.36,1.33]	[-0.48,0.26]	[0.47,2.19]	[-1.53,0.22]	[-0.19,0.35]	[-0.55,3.73]	[-1.44,1.91]	[-0.72,0.27]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.78	-0.06	0.89	-0.18	0.15	0.84	-0.36	0.03
	[0.37,1.20]	[-0.35,0.24]	[0.26,1.51]	[-0.87,0.50]	[-0.05,0.35]	[-0.42,2.10]	[-1.40,0.69]	[-0.37,0.43]
				5 × 5 Size and Short-Term Reversal				
Bayes No Constraint	0.79	0	0.31	0.61	0.21	-0.48	0.32	0.15
	[0.33,1.25]	[-0.35,0.35]	[-0.94,1.57]	[0.04,1.21]	[-0.74,1.17]	[-1.85,0.86]	[-0.87,1.46]	[-0.45,0.76]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]

Table A.3 **continued from previous page**

	MKTRF	SMB	HML	RMW	CMA	UMD	BAB	QMJ
WLS	0.76	0.03	-0.14	0.95	0.15	-0.48	0.31	0.40
	[0.34,1.17]	[-0.27,0.33]	[-1.12,0.83]	[0.44,1.46]	[-0.58,0.88]	[-1.52,0.56]	[-0.63,1.25]	[-0.11,0.91]
	5×5 Size and Long-Term Reversal							
Bayes No Constraint	0.87	0.26	0.02	-0.14	0.09	0.93	1.09	0.01
	[0.38,1.33]	[-0.11,0.63]	[-0.53,0.57]	[-0.63,0.33]	[-0.34,0.52]	[-0.22,2.17]	[-0.02,2.28]	[-0.39,0.44]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.18]	[-0.21,0.36]	[-0.16,0.40]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.85]	[0.38,1.09]	[0.24,0.71]
WLS	0.81	0.22	0.14	-0.16	0.03	1.11	1.08	0.04
	[0.40,1.23]	[-0.09,0.53]	[-0.34,0.63]	[-0.69,0.36]	[-0.38,0.43]	[0.32,1.90]	[0.13,2.03]	[-0.32,0.40]
	All Portfolios							
Bayes No Constraint	0.66	0.1	0.12	0.3	0.2	0.45	0.55	0.24
	[0.20,1.08]	[-0.13,0.30]	[-0.20,0.42]	[0.10,0.50]	[-0.01,0.39]	[-0.02,0.87]	[0.18,0.93]	[0.03,0.44]
Bayes With Constraint	0.72	0.08	0.12	0.4	0.21	0.42	0.73	0.48
	[0.28,1.16]	[-0.21,0.35]	[-0.17,0.41]	[0.19,0.61]	[0.01,0.41]	[-0.01,0.84]	[0.39,1.06]	[0.25,0.71]
WLS	0.77	0.10	0.14	0.31	0.15	0.48	0.45	0.24
	[0.36,1.18]	[-0.18,0.37]	[-0.16,0.44]	[0.06,0.55]	[-0.06,0.36]	[0.06,0.90]	[0.01,0.89]	[-0.03,0.51]

A.III MCMC and Marginal Likelihood: General Case

(MCMC Algorithm) Initialize parameters and iterate

1. **Sample** $\boldsymbol{\beta}_i$ from

$$\boldsymbol{\beta}_i \mid \text{rest} \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}_i, \mathcal{B}_i)$$

where

$$\widehat{\boldsymbol{\beta}}_i = \mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \tilde{\boldsymbol{F}}'\boldsymbol{R}_i/\sigma_i^2$$
$$\mathcal{B}_i^{-1} = \mathcal{B}_{0i}^{-1} + \tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/\sigma_i^2$$
$$\tilde{\boldsymbol{F}}' = (\tilde{\boldsymbol{f}}_1, ..., \tilde{\boldsymbol{f}}_t)'$$

2. **Sample** σ_{ei}^2 from

$$\sigma_{ei}^2 \mid \text{rest} \sim \mathcal{IG}\left(\nu_{ei}, \delta_{ei}\right)$$

where

$$\nu_{ei} = \frac{\nu_{e,0i} + T}{2}$$
$$\delta_e = \frac{\delta_{e,0i} + SSE_i}{2}$$
$$SSE_i = \text{sum}(\boldsymbol{E}_i \odot \boldsymbol{E}_i)$$
$$\boldsymbol{E}_i = \boldsymbol{R}_i - \tilde{\boldsymbol{F}}\boldsymbol{\beta}_i'$$

3. **Sample** $\boldsymbol{\lambda}^{NT}$ from

$$\boldsymbol{\lambda}^{NT} \mid \text{rest} \sim \mathcal{N}(\widehat{\boldsymbol{\lambda}}^{NT}, \boldsymbol{L}^{NT})$$

where

$$(\boldsymbol{L}^{NT})^{-1} = (\boldsymbol{L}_0^{NT})^{-1} + T\boldsymbol{B}^{NT\prime}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{B}^{NT}$$
$$\widehat{\boldsymbol{\lambda}}^{NT} = \boldsymbol{L}^{NT}\left((\boldsymbol{L}_0^{NT})^{-1}\boldsymbol{\lambda}_0^{NT} + \boldsymbol{B}^{NT\prime}\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{R} - \tilde{\boldsymbol{F}}\boldsymbol{B}')\mathbf{1}_T\right)$$

4. **Sample** $\boldsymbol{\mu}$ from

$$\boldsymbol{\mu} \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}, \boldsymbol{D}_1)$$

where

$$\widehat{\boldsymbol{\mu}} = \boldsymbol{D}_1(\boldsymbol{M}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_u^{-1}\boldsymbol{F}'\mathbf{1}_T)$$
$$\boldsymbol{D}_1^{-1} = \boldsymbol{M}_1^{-1} + T\boldsymbol{\Sigma}_u^{-1}$$

and

$$\boldsymbol{\mu}_1 = \boldsymbol{M}_1(\boldsymbol{M}_0^{-1}\boldsymbol{\mu}_0 + \tilde{\boldsymbol{B}}'\boldsymbol{\Sigma}_e^{-1}\boldsymbol{R}_*)$$

$$\boldsymbol{M}_1^{-1} = \boldsymbol{M}_0^{-1} + T\tilde{\boldsymbol{B}}'\boldsymbol{\Sigma}_e^{-1}\tilde{\boldsymbol{B}}$$

$$\boldsymbol{R}_* = colsum(\boldsymbol{R} - \boldsymbol{F}\boldsymbol{B}' - \boldsymbol{1}'\boldsymbol{\lambda}^{NT'}\boldsymbol{B}^{NT'})'$$

$\tilde{\boldsymbol{B}}$ is an $n \times k^*$ matrix where the corresponding $\boldsymbol{f}^{NT}$ columns contain the $\boldsymbol{\beta}$ for $\boldsymbol{f}^{NT}$ and leave the other columns zero.

5. **Sample** $\boldsymbol{\Sigma}_u$ from

$$\boldsymbol{\Sigma}_u \mid \text{rest} \sim \mathcal{IW}\left(\nu_u, \boldsymbol{S}_u\right)$$

where

$$\nu_u = \rho_{u,0} + T$$

$$\boldsymbol{S}_u = \boldsymbol{Q}_{u,0} + \boldsymbol{Q}$$

$$\boldsymbol{Q} = (\boldsymbol{F} - \boldsymbol{1}_T\boldsymbol{\mu}')'(\boldsymbol{F} - \boldsymbol{1}_T\boldsymbol{\mu}')$$

Algorithm MCMC Sampling with Parameter Expressions

1: **Initialize parameters**: Set $\boldsymbol{\theta}^{(0)} = (\boldsymbol{B}^{(0)}, \boldsymbol{\Sigma}_e^{(0)}, \boldsymbol{\lambda}^{NT(0)}, \boldsymbol{\Sigma}_u^{(0)}, \boldsymbol{\mu}^{(0)})$
2: **for** $g = 1$ to $n_0 + M$ **do**
3: **Step 1:** For $i = 1$ to $i = n$, sample $\boldsymbol{\beta}_i^{(g+1)} \mid \boldsymbol{\Sigma}_e^{(g)}, \boldsymbol{\lambda}^{NT(g)}, \boldsymbol{\Sigma}_u^{(g)}, \boldsymbol{\mu}^{(g)} \sim \mathcal{N}\left(\widehat{\boldsymbol{\beta}}_i^{(g)}, \mathcal{B}_i^{(g)}\right)$
4: **Step 2:** For $i = 1$ to $i = n$, sample $\sigma_{ei}^2 \mid \boldsymbol{B}^{(g+1)}, \boldsymbol{\lambda}^{NT(g)}, \boldsymbol{\Sigma}_u^{(g)}, \boldsymbol{\mu}^{(g)} \sim \mathcal{IG}\left(\nu_{ei}^{(g)}, \delta_{ei}^{(g)}\right)$
5: **Step 3:** Sample $\boldsymbol{\lambda}^{NT} \mid \boldsymbol{B}^{(g+1)}, \boldsymbol{\Sigma}_e^{(g+1)}, \boldsymbol{\Sigma}_u^{(g)}, \boldsymbol{\mu}^{(g)} \sim \mathcal{N}\left(\widehat{\boldsymbol{\lambda}}^{NT(g)}, \boldsymbol{L}^{NT(g)}\right)$
6: **Step 4:** Sample $\boldsymbol{\mu}^{(g+1)} \mid \boldsymbol{B}^{(g+1)}, \boldsymbol{\Sigma}_e^{(g+1)}, \boldsymbol{\lambda}^{NT(g+1)}, \boldsymbol{\Sigma}_u^{(g)} \sim \mathcal{N}\left(\widehat{\boldsymbol{\mu}}^{(g)}, \boldsymbol{D}_1^{(g)}\right)$
7: **Step 5:** Sample $\boldsymbol{\Sigma}_u^{(g+1)} \mid \boldsymbol{B}^{(g+1)}, \boldsymbol{\Sigma}_e^{(g+1)}, \boldsymbol{\lambda}^{NT(g+1)}, \boldsymbol{\mu}^{(g+1)} \sim \mathcal{IW}\left(\nu_u^{(g)}, \boldsymbol{S}_u^{(g)}\right)$
8: **end for**
9: After burn-in, use samples $\{\boldsymbol{\theta}^{(g)}\}_{g=1}^{N}$ to estimate the posterior distribution.

(Marginal Likelihood) To estimate the marginal likelihood of each contending model, we employ the Chib (1995) method, which starts with the convenient expression of the log-marginal likelihood

$$\ln m(\boldsymbol{y}_{1:T} \mid \mathcal{M}_j) = \ln \pi(\boldsymbol{\theta}^* \mid \mathcal{M}_j) + \ln p(\boldsymbol{y}_{1:T} \mid \mathcal{M}_j, \boldsymbol{\theta}^*)$$

$$- \ln \pi(\boldsymbol{\theta}^* \mid \mathcal{M}_j, \boldsymbol{y}_{1:T})$$

where $\boldsymbol{\theta}^* = (\boldsymbol{B}^*, \boldsymbol{\Sigma}_e^*, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\mu}^*)$ is some chosen point, say the posterior mean. In this expression, the prior and likelihood ordinates can be found analytically. As for the third term, the posterior ordinate suppresses the model index and uses a marginal-conditional decomposition to write

$$\ln \pi(\boldsymbol{\theta}^* \mid \boldsymbol{y}_{1:T}) = \ln \pi(\boldsymbol{B}^* \mid \boldsymbol{y}_{1:T}) + \ln \pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^*)$$
$$+ \ln \pi(\boldsymbol{\lambda}^{NT*} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*) + \ln \pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*)$$
$$+ \ln \pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*)$$

An estimate of the first ordinate is obtained as

$$\pi(\boldsymbol{B}^* \mid \boldsymbol{y}_{1:T}) = \frac{1}{G} \sum_{g=1}^{G} \prod_{i=1}^{n} \mathcal{N}\left(\mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \tilde{\boldsymbol{F}}^{\prime(g)}\boldsymbol{R}_i/\sigma_i^{2(g)}, (\mathcal{B}_{0i}^{-1} + \tilde{\boldsymbol{F}}^{\prime(g)}\tilde{\boldsymbol{F}}^{(g)}/\sigma_i^{2(g)})^{-1}\right)$$

the second from reduced run 1

$$\pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^*) = \frac{1}{G} \sum_{g=1}^{G} \prod_{i=1}^{n} \mathcal{IG}\left(\frac{\nu_{e,0i} + T}{2}, \frac{\delta_{e,0i} + \text{sum}((\boldsymbol{R}_i - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{\beta}_i^{*\prime}) \odot (\boldsymbol{R}_i - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{\beta}_i^{*\prime}))}{2}\right)$$

the third from reduced run 2

$$\pi(\boldsymbol{\lambda}^{NT*} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*) = \frac{1}{G} \sum_{g=1}^{G} \mathcal{N}\Bigg(\boldsymbol{L}^{NT}\Big((\boldsymbol{L}_0^{NT})^{-1}\boldsymbol{\lambda}_0^{NT} + \boldsymbol{B}^{*\prime NT}\boldsymbol{\Sigma}_e^{*-1}(\boldsymbol{R} - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{B}^{*\prime})\mathbf{1}_T\Big),$$
$$\Big((\boldsymbol{L}_0^{NT})^{-1} + T\boldsymbol{B}^{*\prime NT}\boldsymbol{\Sigma}_e^{*-1}\boldsymbol{B}^{*NT}\Big)^{-1}\Bigg)$$

the fourth from reduced run 3

$$\pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*) = \mathcal{IW}\left(\rho_{u,0} + T, \boldsymbol{Q}_{u,0} + (\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}^{(g)\prime})^{\prime}(\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}^{(g)\prime})\right)$$

finally, the fifth from the chosen point

$$\pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^*) = \mathcal{N}\left(\widehat{\boldsymbol{\mu}}^*, \boldsymbol{D}_1^*\right)$$

where

$$\widehat{\boldsymbol{\mu}}^* = \boldsymbol{D}_1(\boldsymbol{M}_1^{*-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_u^{*-1}\boldsymbol{F}^{\prime}\mathbf{1}_T)$$
$$\boldsymbol{D}_1^{-1} = \boldsymbol{M}_1^{*-1} + T\boldsymbol{\Sigma}_u^{*-1}$$

and

$$\boldsymbol{\mu}_1 = \boldsymbol{M}_1^*(\boldsymbol{M}_0^{-1}\boldsymbol{\mu}_0 + \tilde{\boldsymbol{B}}^{*\prime}\boldsymbol{\Sigma}_e^{*-1}\boldsymbol{R}_*)$$
$$\boldsymbol{M}_1^{*-1} = \boldsymbol{M}_0^{-1} + T\tilde{\boldsymbol{B}}^{*\prime}\boldsymbol{\Sigma}_e^{*-1}\tilde{\boldsymbol{B}}^*$$
$$\boldsymbol{R}_* = colsum(\boldsymbol{R} - \boldsymbol{F}\boldsymbol{B}^{*\prime} - \mathbf{1}^{\prime}\boldsymbol{\lambda}^{NT*\prime}\boldsymbol{B}^{*NT\prime})^{\prime}$$

A.IV MCMC and Marginal Likelihood: NT Factors Only

In this section, we propose the model specification with only nontraded factors, such as a consumption-based model. Start with the restricted TS model, which is the

same as equation 7

$$\boldsymbol{r}_t = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t \tag{A.1}$$

and factors

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t \tag{A.2}$$

The observed data is $\boldsymbol{y}_{1:T} = \{\boldsymbol{r}_{1:T}, \boldsymbol{f}^*_{1:T}\}$, excluding the training sample. We first specify the conditional models that are used to construct the posterior distributions involved in the MCMC sampling. For updating $\boldsymbol{B}^{NT} = [\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_n]$ and $\boldsymbol{\Sigma}_e$ we use the model

$$\boldsymbol{r}_t = \boldsymbol{B}^{NT}\tilde{\boldsymbol{f}}_t^{NT} + \boldsymbol{e}_t$$

where

$$\tilde{\boldsymbol{f}}_t^{NT} = \boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT} + \boldsymbol{\lambda}^{NT} \tag{A.3}$$

For updating $\boldsymbol{\Sigma}_u$ we use the model

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t$$

For updating $\boldsymbol{\lambda}^{NT}$ we use the model

$$\boldsymbol{r}_t - \boldsymbol{B}^{NT}\boldsymbol{f}_t^{NT} + \boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} = \boldsymbol{\lambda}^{NT} + \boldsymbol{e}_t$$

Finally, for updating $\boldsymbol{\mu}$ we use the models

$$\boldsymbol{r}_t - \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} - \boldsymbol{B}^{NT}\boldsymbol{f}_t^{NT} = -\boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} + \boldsymbol{e}_t$$

and

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t$$

(MCMC Algorithm: NT Factors Only) Initialize parameters and iterate

1. **Sample** $\boldsymbol{\beta}_i$ from

$$\boldsymbol{\beta}_i \mid \text{rest} \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}_i, \mathcal{B}_i)$$

where

$$\widehat{\boldsymbol{\beta}}_i = \mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \tilde{\boldsymbol{F}}' R_i/\sigma_i^2$$
$$\mathcal{B}_i^{-1} = \mathcal{B}_{0i}^{-1} + \tilde{\boldsymbol{F}}'\tilde{\boldsymbol{F}}/\sigma_i^2$$
$$\tilde{\boldsymbol{F}}' = (\tilde{\boldsymbol{f}}_1^{NT}, ..., \tilde{\boldsymbol{f}}_t^{NT})'$$

2. **Sample** σ_{ei}^2 from

$$\sigma_{ei}^2 \mid \text{rest} \sim \mathcal{IG}(\nu_{ei}, \delta_{ei})$$

where

$$\nu_{ei} = \frac{\nu_{e,0i} + T}{2}$$
$$\delta_e = \frac{\delta_{e,0i} + SSE_i}{2}$$
$$SSE_i = \text{sum}(\boldsymbol{E}_i \odot \boldsymbol{E}_i)$$
$$\boldsymbol{E}_i = \boldsymbol{R}_i - \tilde{\boldsymbol{F}}\boldsymbol{\beta}_i'$$

3. **Sample $\boldsymbol{\lambda}$** from

$$\boldsymbol{\lambda} \mid \text{rest} \sim \mathcal{N}(\widehat{\boldsymbol{\lambda}}, \boldsymbol{L})$$

where

$$\boldsymbol{L}^{-1} = (\boldsymbol{L}_0)^{-1} + T\boldsymbol{B}'^{NT}\boldsymbol{\Sigma}_e^{-1}\boldsymbol{B}^{NT}$$
$$\widehat{\boldsymbol{\lambda}^{NT}} = \boldsymbol{L}\left((\boldsymbol{L}_0)^{-1}\boldsymbol{\lambda}_0 + \boldsymbol{B}'^{NT}\boldsymbol{\Sigma}_e^{-1}(\boldsymbol{R} - \tilde{\boldsymbol{F}}\boldsymbol{B}'^{NT})\mathbf{1}_T\right)$$

4. **Sample $\boldsymbol{\mu}$** from

$$\boldsymbol{\mu} \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}, \boldsymbol{D}_1)$$

where

$$\widehat{\boldsymbol{\mu}} = \boldsymbol{D}_1(\boldsymbol{M}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_u^{-1}\boldsymbol{F}'\mathbf{1}_T)$$
$$\boldsymbol{D}_1^{-1} = \boldsymbol{M}_1^{-1} + T\boldsymbol{\Sigma}_u^{-1}$$

and

$$\boldsymbol{\mu}_1 = \boldsymbol{M}_1(\boldsymbol{M}_0^{-1}\boldsymbol{\mu}_0 + \tilde{\boldsymbol{B}}'\boldsymbol{\Sigma}_e^{-1}\boldsymbol{R}_*)$$
$$\boldsymbol{M}_1^{-1} = \boldsymbol{M}_0^{-1} + T\tilde{\boldsymbol{B}}'\boldsymbol{\Sigma}_e^{-1}\tilde{\boldsymbol{B}}$$
$$\boldsymbol{R}_* = colsum(\boldsymbol{R} - \boldsymbol{F}^{NT}\boldsymbol{B}'^{NT} - \mathbf{1}'\boldsymbol{\lambda}'^{NT}\boldsymbol{B}'^{NT})'$$

$\tilde{\boldsymbol{B}}$ is an $n \times k^*$ matrix where the corresponding f columns contain the $\boldsymbol{\beta}$ for f and leave the other columns zero.

5. **Sample $\boldsymbol{\Sigma}_u$** from

$$\boldsymbol{\Sigma}_u \mid \text{rest} \sim \mathcal{IW}(\nu_u, \boldsymbol{S}_u)$$

where

$$\nu_u = \rho_{u,0} + T$$
$$\boldsymbol{S}_u = \boldsymbol{Q}_{u,0} + \boldsymbol{Q}$$
$$\boldsymbol{Q} = (\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}')'(\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}')$$

The marginal likelihood estimation is almost the same as the case of both traded

and nontraded factors. Still, the only difference is that $\tilde{\boldsymbol{F}}$ is not a combination of traded and nontraded factors, but the adjusted nontraded factor in equation A.3.

(Marginal Likelihood: NT Factors Only) To estimate the marginal likelihood of each contending model, we again employ the Chib (1995) method, where now $\boldsymbol{\theta} = (\boldsymbol{B}^{NT}, \boldsymbol{\Sigma}_e, \boldsymbol{\lambda}^{NT}, \boldsymbol{\Sigma}_u, \boldsymbol{\mu})$. The posterior ordinate is obtained via the following marginal-conditional decomposition:

$$\begin{aligned}\ln \pi(\boldsymbol{\theta}^* \mid \boldsymbol{y}_{1:T}) = \ln \pi(\boldsymbol{B}^{*NT} \mid \boldsymbol{y}_{1:T}) + \ln \pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^{*NT}) \\ + \ln \pi(\boldsymbol{\lambda}^{NT*} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}) + \ln \pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}) \\ + \ln \pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}).\end{aligned}$$

The first ordinate is estimated as:

$$\begin{aligned}\pi(\boldsymbol{B}^{*NT} \mid \boldsymbol{y}_{1:T}) = \frac{1}{G}\sum_{g=1}^{G}\prod_{i=1}^{n}\mathcal{N}\Big(\mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \tilde{\boldsymbol{F}}'^{(g)}R_i/\sigma_i^{2(g)}, \\ (\mathcal{B}_{0i}^{-1} + \tilde{\boldsymbol{F}}'^{(g)}\tilde{\boldsymbol{F}}^{(g)}/\sigma_i^{2(g)})^{-1}\Big).\end{aligned}$$

The second is estimated from reduced run 1:

$$\begin{aligned}\pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^{*NT}) = \frac{1}{G}\sum_{g=1}^{G}\prod_{i=1}^{n}\mathcal{IG}\left(\frac{\nu_{e,0i}+T}{2}, \right. \\ \left. \frac{\delta_{e,0i} + \text{sum}((\boldsymbol{R}_i - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{\beta}_i^{*\prime}) \odot (\boldsymbol{R}_i - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{\beta}_i^{*\prime}))}{2}\right).\end{aligned}$$

The third from reduced run 2:

$$\begin{aligned}\pi(\boldsymbol{\lambda}^{NT*} \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}) = \frac{1}{G}\sum_{g=1}^{G}\mathcal{N}\Bigg(\boldsymbol{L}^{NT}\Big((\boldsymbol{L}_0^{NT})^{-1}\boldsymbol{\lambda}_0^{NT} + \boldsymbol{B}^{*\prime NT}\boldsymbol{\Sigma}_e^{*-1} \\ \cdot (\boldsymbol{R} - \tilde{\boldsymbol{F}}^{(g)}\boldsymbol{B}^{*\prime NT})\mathbf{1}_T\Big), \Big((\boldsymbol{L}_0^{NT})^{-1} + T\boldsymbol{B}^{*\prime NT}\boldsymbol{\Sigma}_e^{*-1}\boldsymbol{B}^{*NT}\Big)^{-1}\Bigg).\end{aligned}$$

The fourth from reduced run 3:

$$\begin{aligned}\pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}) = \mathcal{IW}\left(\rho_{u,0} + T, \right. \\ \left. \boldsymbol{Q}_{u,0} + (\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}^{(g)\prime})'(\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}^{(g)\prime})\right).\end{aligned}$$

Finally, the fifth is from the chosen point:

$$\pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\lambda}^{NT*}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*NT}) = \mathcal{N}\left(\widehat{\boldsymbol{\mu}}^*, \boldsymbol{D}_1^*\right),$$

where:

$$\widehat{\boldsymbol{\mu}}^* = \boldsymbol{D}_1(\boldsymbol{M}_1^{*-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_u^{*-1}\boldsymbol{F}'\mathbf{1}_T), \quad \boldsymbol{D}_1^{-1} = \boldsymbol{M}_1^{*-1} + T\boldsymbol{\Sigma}_u^{*-1}.$$

and

$$\boldsymbol{\mu}_1 = \boldsymbol{M}_1^*(\boldsymbol{M}_0^{-1}\boldsymbol{\mu}_0 + \tilde{\boldsymbol{B}}^{*\prime}\boldsymbol{\Sigma}_e^{*-1}\boldsymbol{R}_*), \quad \boldsymbol{M}_1^{*-1} = \boldsymbol{M}_0^{-1} + T\tilde{\boldsymbol{B}}^{*\prime}\boldsymbol{\Sigma}_e^{*-1}\tilde{\boldsymbol{B}}^*.$$

Finally, we define:

$$\boldsymbol{R}_* = \text{colsum}(\boldsymbol{R} - \boldsymbol{F}\boldsymbol{B}^{*NT\prime} - \mathbf{1}'\boldsymbol{\lambda}^{NT*\prime}\boldsymbol{B}^{*NT\prime})'.$$

A.V MCMC and Marginal Likelihood: T Factors Only

In this section, we propose the model specification with only traded factors, for example, the Fama-French 3-factor model. In this case, the restricted TS model takes a different form without λ

$$\boldsymbol{r}_t = \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{e}_t \tag{A.4}$$

and factors

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t \tag{A.5}$$

The observed data is $\boldsymbol{y}_{1:T} = \{r_{1:T}, f_{1:T}^*\}$ denotes the sample data exclude the training sample. We first specify the conditional model used to sample these posterior distributions. For updating $\boldsymbol{B}^T = [\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_n]$ and $\boldsymbol{\Sigma}_e$ we use the model

$$\boldsymbol{r}_t = \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{e}_t$$

For updating $\boldsymbol{\Sigma}_u$ we use the model

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t$$

Finally, for updating $\boldsymbol{\mu}$ we use the models

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t$$

(MCMC Algorithm with Only Traded Factors) Initialize parameters and iterate

1. **Sample** $\boldsymbol{\beta}_i$ from

$$\boldsymbol{\beta}_i \mid \text{rest} \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}_i, \mathcal{B}_i)$$

where

$$\widehat{\boldsymbol{\beta}}_i = \mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \boldsymbol{F}'R_i/\sigma_i^2$$
$$\mathcal{B}_i^{-1} = \mathcal{B}_{0i}^{-1} + \boldsymbol{F}'\boldsymbol{F}/\sigma_i^2$$
$$\boldsymbol{F}' = (\boldsymbol{f}_1^T, ..., \boldsymbol{f}_t^T)'$$

2. **Sample** σ_{ei}^2 from

$$\sigma_{ei}^2 \mid \text{rest} \sim \mathcal{IG}(\nu_{ei}, \delta_{ei})$$

where

$$\nu_{ei} = \frac{\nu_{e,0i} + T}{2}$$
$$\delta_e = \frac{\delta_{e,0i} + SSE_i}{2}$$
$$SSE_i = \text{sum}(\boldsymbol{E}_i \odot \boldsymbol{E}_i)$$
$$\boldsymbol{E}_i = \boldsymbol{R}_i - \boldsymbol{F}\boldsymbol{\beta}_i'$$

3. **Sample $\boldsymbol{\mu}$** from

$$\boldsymbol{\mu} \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}, \boldsymbol{D}_1)$$

where

$$\widehat{\boldsymbol{\mu}} = \boldsymbol{D}_1(\boldsymbol{M}_0^{-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_u^{-1}\boldsymbol{F}'\mathbf{1}_T)$$
$$\boldsymbol{D}_1^{-1} = \boldsymbol{M}_0^{-1} + T\boldsymbol{\Sigma}_u^{-1}$$

4. **Sample $\boldsymbol{\Sigma}_u$** from

$$\boldsymbol{\Sigma}_u \mid \text{rest} \sim \mathcal{IW}(\nu_u, \boldsymbol{S}_u)$$

where

$$\nu_u = \rho_{u,0} + T$$
$$S_u = \boldsymbol{Q}_{u,0} + \boldsymbol{Q}$$
$$\boldsymbol{Q} = (\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}')'(\boldsymbol{F} - \mathbf{1}_T\boldsymbol{\mu}')$$

(Marginal Likelihood: T Factors Only) To estimate the marginal likelihood of each contending model, we again employ the Chib (1995) method where now $\boldsymbol{\theta} = (\boldsymbol{B}^T, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_u, \boldsymbol{\mu})$. The posterior ordinate is obtained via the following marginal-conditional decomposition:

$$\ln \pi(\boldsymbol{\theta}^* \mid \boldsymbol{y}_{1:T}) = \ln \pi(\boldsymbol{B}^{*T} \mid \boldsymbol{y}_{1:T}) + \ln \pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^{*T})$$
$$+ \ln \pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*T}) + \ln \pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*T})$$

An estimate of the first ordinate is obtained as

$$\pi(\boldsymbol{B}^{*T} \mid \boldsymbol{y}_{1:T}) = \frac{1}{G}\sum_{g=1}^{G}\prod_{i=1}^{n}\mathcal{N}\Big(\mathcal{B}_{0i}\boldsymbol{\beta}_{0i} + \boldsymbol{F}'^{(g)}R_i/\sigma_i^{2(g)},$$
$$(\mathcal{B}_{0i}^{-1} + \boldsymbol{F}'^{(g)}\boldsymbol{F}^{(g)}/\sigma_i^{2(g)})^{-1}\Big)$$

the second from reduced run 1

$$\pi(\boldsymbol{\Sigma}_e^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{B}^{*T}) = \frac{1}{G}\sum_{g=1}^{G}\prod_{i=1}^{n}\mathcal{IG}\left(\frac{\nu_{e,0i}+T}{2}, \frac{\delta_{e,0i}+\text{sum}((\boldsymbol{R}_i-\boldsymbol{F}^{(g)}\boldsymbol{\beta}_i^{*\prime})\odot(\boldsymbol{R}_i-\boldsymbol{F}^{(g)}\boldsymbol{\beta}_i^{*\prime}))}{2}\right)$$

the third from reduced run 2

$$\pi(\boldsymbol{\Sigma}_u^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*T}) = \mathcal{IW}\left(\rho_{u,0}+T, \boldsymbol{Q}_{u,0}+(\boldsymbol{F}-\boldsymbol{1}_T\boldsymbol{\mu}^{(g)'})'(\boldsymbol{F}-\boldsymbol{1}_T\boldsymbol{\mu}^{(g)'})\right)$$

finally, the fourth from the chosen point

$$\pi(\boldsymbol{\mu}^* \mid \boldsymbol{y}_{1:T}, \boldsymbol{\Sigma}_u^*, \boldsymbol{\Sigma}_e^*, \boldsymbol{B}^{*T}) = \mathcal{N}(\widehat{\boldsymbol{\mu}}^*, \boldsymbol{D}_1^*)$$

where

$$\widehat{\boldsymbol{\mu}}^* = \boldsymbol{D}_1(\boldsymbol{M}_0^{*-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_u^{*-1}\boldsymbol{F}'\boldsymbol{1}_T)$$
$$\boldsymbol{D}_1^{-1} = \boldsymbol{M}_0^{*-1} + T\boldsymbol{\Sigma}_u^{*-1}$$

A.VI Model Specification with Zero-Beta Rate

In this section, we propose the model specification with zero beta rate under the general setup. The return model in equation 4 is augmented with a zero beta rate λ_0,

$$\boldsymbol{r}_t = \boldsymbol{1}\lambda_0 + \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}) + \boldsymbol{e}_t \tag{A.6}$$

and factors

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t \tag{A.7}$$

The posterior distribution for the parameters of the model, given data from $t = 1, \ldots, T$, is proportional to the product of the likelihood and the prior. The return model likelihood is

$$p(\boldsymbol{r}_t \mid \lambda_0, \boldsymbol{f}_t^*, \boldsymbol{B}^T, \boldsymbol{B}^{NT}, \boldsymbol{\lambda}^{NT}, \boldsymbol{\Sigma}_e) = \prod_{t=1}^{T}\mathcal{N}(\boldsymbol{r}_t \mid \boldsymbol{1}\lambda_0 + \boldsymbol{B}^T\boldsymbol{f}_t^T + \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{B}^{NT}(\boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT}), \boldsymbol{\Sigma}_e),$$

where $\boldsymbol{\Sigma}_e = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, and the factor model likelihood is

$$p(\boldsymbol{f}_t^* \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}_u) = \prod_{t=1}^{T}\mathcal{N}(\boldsymbol{f}_t^* \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}_u).$$

The posterior distribution is proportional to the product of the two likelihoods and the prior on the parameters:

$$p(\lambda_0, \boldsymbol{B}^T, \boldsymbol{B}^{NT}, \boldsymbol{\lambda}^{NT}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_u).$$

We sample the resulting posterior distribution using MCMC methods. In this

sampling, we consider the conditional distributions of blocks of parameters conditioned on the remaining parameters and sample these conditional distributions in sequence to produce a Markov chain whose invariant distribution is the posterior distribution (Chib, 2001). The conditional distributions of these different blocks of parameters are derived from the following formulations of the joint model. For updating $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_e$ we write the joint model as

$$\boldsymbol{r}_t = \mathbf{1}\lambda_0 + \boldsymbol{B}\tilde{\boldsymbol{f}}_t + \boldsymbol{e}_t$$

where

$$\tilde{\boldsymbol{f}}_t = \begin{pmatrix} \boldsymbol{f}_t^T \\ \boldsymbol{f}_t^{NT} - \boldsymbol{\mu}^{NT} + \boldsymbol{\lambda}^{NT} \end{pmatrix}$$

from which the conditional posterior distributions of $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_e$ can be derived by Bayesian results for multivariate regression models. The updating of $\boldsymbol{\Sigma}_u$ arises easily from the factor model. For updating $\boldsymbol{\lambda}^{NT}$ we rewrite the full model now as

$$\boldsymbol{r}_t - \mathbf{1}\lambda_0 - \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} = \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} + \boldsymbol{e}_t$$

For updating λ_0 we rewrite the full model as

$$\boldsymbol{r}_t - \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} - \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} = \mathbf{1}\lambda_0 + \boldsymbol{e}_t$$

Finally, for updating $\boldsymbol{\mu}$ we write the joint model as

$$\boldsymbol{r}_t - \mathbf{1}\lambda_0 - \boldsymbol{B}^{NT}\boldsymbol{\lambda}^{NT} - \boldsymbol{B}\boldsymbol{f}_t = -\boldsymbol{B}^{NT}\boldsymbol{\mu}^{NT} + \boldsymbol{e}_t$$

and

$$\boldsymbol{f}_t^* = \boldsymbol{\mu} + \boldsymbol{u}_t$$

and then derive the conditional posterior of $\boldsymbol{\mu}$ in two steps, first with the return model likelihood and then with the factor model likelihood. The resulting MCMC algorithm is similar to previous cases, except that there is an additional block for sampling λ_0 from

$$\lambda_0 \sim \mathcal{N}(\widehat{\mu}_{\lambda_0}, \widehat{\sigma}^2_{\lambda_0}),$$

where, on letting $\tilde{\boldsymbol{R}} = \boldsymbol{R} - \tilde{\boldsymbol{F}}\boldsymbol{B} - \mathbf{1}'\boldsymbol{\lambda}^{NT'}\boldsymbol{B}^{NT'}$,

$$\widehat{\mu}_{\lambda_0} = \widehat{\sigma}^2_{\lambda_0}(\mathbf{1}'\boldsymbol{\Sigma}_e^{-1}\tilde{\boldsymbol{R}} + \widehat{\mu}_{0\lambda_0}\sigma_{0\lambda_0}^{-2}), \quad \widehat{\sigma}^{-2}_{\lambda_0} = \sigma_{0\lambda_0}^{-2} + \mathbf{1}'\boldsymbol{\Sigma}_e^{-1}\mathbf{1} \quad .$$

References

Allena, R. (2020). Comparing asset pricing models with non-traded factors and principal components. *Available at SSRN 3669940*.

Asness, C. S., A. Frazzini, and L. H. Pedersen (2019). Quality minus junk. *Review of Accounting Studies 24*(1), 34–112.

Avramov, D. and J. C. Chao (2006). An exact Bayes test of asset pricing models with application to international markets. *Journal of Business 79*(1), 293–324.

Barillas, F., R. Kan, C. Robotti, and J. Shanken (2020). Model comparison with Sharpe ratios. *Journal of Financial and Quantitative Analysis 55*(6), 1840–1874.

Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *Journal of Finance 73*(2), 715–754.

Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger (1989). Empirical tests of the consumption-oriented CAPM. *Journal of Finance 44*(2), 231–262.

Bryzgalova, S., J. Huang, and C. Julliard (2023). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *Journal of Finance 78*(1), 487–557.

Chan, K., S. Foresi, and L. H. Lang (1996). Does money explain asset returns? theory and empirical analysis. *Journal of Finance 51*(1), 345–361.

Chen, N.-F., R. Roll, and S. A. Ross (1986). Economic forces and the stock market. *Journal of Business*, 383–403.

Chen, Z., Z. Da, D. Huang, and L. Wang (2023). Presidential economic approval rating and the cross-section of stock returns. *Journal of Financial Economics 147*(1), 106–131.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association 90*(432), 1313–1321.

Chib, S. (2001). Markov Chain Monte Carlo methods: Computation and inference. *Handbook of Econometrics 5*, 3569–3649.

Chib, S., M. Shin, and A. Simoni (2018). Bayesian estimation and comparison of moment condition models. *Journal of the American Statistical Association 113*(524), 1656–1668.

Chib, S. and X. Zeng (2020). Which factors are risk factors in asset pricing? A model scan framework. *Journal of Business & Economic Statistics 38*, 771–783.

Chib, S., X. Zeng, and L. Zhao (2020). On comparing asset pricing models. *Journal of Finance 75*(1), 551–577.

Chib, S., L. Zhao, and G. Zhou (2024). Winners from winners: A tale of tisk tactors. *Management Science 70*(1), 396–414.

Cochrane, J. H. (2009). *Asset pricing: Revised edition*. Princeton university press.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fama, E. F. and K. R. French (2018). Choosing factors. *Journal of Financial Economics 128*(2), 234–252.

Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo: A test of new factors. *Journal of Finance 75*(3), 1327–1370.

Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics 111*(1), 1–25.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica*, 1121–1152.

Giglio, S. and D. Xiu (2021). Asset pricing with omitted factors. *Journal of Political Economy 129*(7), 1947–1990.

Giglio, S., D. Xiu, and D. Zhang (2024). Test assets and weak factors. *Journal of Finance, Forthcoming*.

Gospodinov, N., R. Kan, and C. Robotti (2014). Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *Review of Financial Studies 27*(7), 2139–2170.

Gospodinov, N., R. Kan, and C. Robotti (2019). Too good to be true? fallacies in evaluating risk factor models. *Journal of Financial Economics 132*(2), 451–471.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 1029–1054.

Hansen, L. P. and R. Jagannathan (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance 52*(2), 557–590.

He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics 126*(1), 1–35.

Hou, K., H. Mo, C. Xue, and L. Zhang (2021). An augmented q-factor model with expected growth. *Review of Finance 25*(1), 1–41.

Jagannathan, R. and Y. Wang (2007). Lazy investors, discretionary consumption, and the cross-section of stock returns. *Journal of Finance 62*(4), 1623–1661.

Jagannathan, R. and Z. Wang (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance 51*(1), 3–53.

Kan, R. and C. Robotti (2009). Model comparison using the Hansen-Jagannathan distance. *Review of Financial Studies 22*(9), 3449–3490.

Kan, R., C. Robotti, and J. Shanken (2013). Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance 68*(6), 2617–2649.

Kan, R., X. Wang, and X. Zheng (2024). In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models. *Journal of Financial Economics 155*, 103837.

Kan, R. and C. Zhang (1999). Two-pass tests of asset pricing models with useless factors. *Journal of Finance 54*(1), 203–235.

Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics 149*(2), 149–173.

Kleibergen, F. and Z. Zhan (2020). Robust inference for consumption-based asset pricing. *Journal of Finance 75*(1), 507–550.

Kroencke, T. A. (2017). Asset pricing without garbage. *Journal of Finance 72*(1), 47–98.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance 56*(3), 815–849.

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics 96*(2), 175–194.

Parker, J. A. and C. Julliard (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy 113*(1), 185–222.

Pastor, L. and R. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Petkova, R. (2006). Do the Fama–French factors proxy for innovations in predictive variables? *Journal of Finance 61*(2), 581–612.

Savov, A. (2011). Asset pricing with garbage. *Journal of Finance 66*(1), 177–201.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance 19*(3), 425–442.

Sweeney, R. J. and A. D. Warga (1986). The pricing of interest-rate risk: evidence from the stock market. *Journal of Finance 41*(2), 393–410.

Yogo, M. (2006). A consumption-based explanation of expected stock returns. *Journal of Finance 61*(2), 539–580.