PHBS WORKING PAPER SERIES

Survival of the Fittest: A Three-Factor Model in the Currency Market

Sining Liu
Soochow University

Yan Wang
City University of Macau

Lingxiao Zhao
Peking University

January 2026

Working Paper 20260104

Abstract

Recent decades have witnessed a rapid growth in proposed risk factors in the international currency market. While most currency factors are evaluated using conventional asset pricing tests that can be mechanically distorted, we identify the best asset pricing model using a survival-of-the-fittest approach. Our framework treats competing currency factor models as alternative specifications of the stochastic discount factor and compares them within a Bayesian framework. We find that a parsimonious three-factor model including the dollar (DOL), carry (CAR), and business-cycle or output-gap (GAP) factors dominates the model space. This model achieves the highest marginal likelihoods, delivers positive and precisely estimated risk premia, and spans the remaining factors. Its dominance is robust across distributional assumptions, alternative factor proxies, and out-of-sample tests. Economically, the model also delivers strong pricing and investment performance, spanning most external currency portfolios and providing superior risk–return trade-offs. Overall, our results provide an economically interpretable benchmark for understanding risk compensation in foreign exchange markets.

Keywords: currency markets, asset pricing, factor models, bayesian model comparison, heavy-tailed returns
JEL Classification: F31, G15, G12, C11

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Survival of the Fittest:

A Three-Factor Model in the Currency Market*

Sining Liu† Yan Wang‡ Lingxiao Zhao§

Abstract

Recent decades have witnessed a rapid growth in proposed risk factors in the international currency market. While most currency factors are evaluated using conventional asset pricing tests that can be mechanically distorted, we identify the best asset pricing model using a survival-of-the-fittest approach. Our framework treats competing currency factor models as alternative specifications of the stochastic discount factor and compares them within a Bayesian framework. We find that a parsimonious three-factor model including the dollar (DOL), carry (CAR), and business-cycle or output-gap (GAP) factors dominates the model space. This model achieves the highest marginal likelihoods, delivers positive and precisely estimated risk premia, and spans the remaining factors. Its dominance is robust across distributional assumptions, alternative factor proxies, and out-of-sample tests. Economically, the model also delivers strong pricing and investment performance, spanning most external currency portfolios and providing superior risk–return trade-offs. Overall, our results provide an economically interpretable benchmark for understanding risk compensation in foreign exchange markets.

Keywords: Currency markets; Asset pricing; Factor models; Bayesian model comparison; Heavy-tailed returns

JEL Classification: F31, G15, G12, C11

*This project is supported by the National Natural Science Foundation of China (Grant No. 72503015), the Shenzhen Outstanding Talent Program for Scientific and Technological Innovation, Doctoral Basic Research Startup Grant (Grant No. RCBS20231211090520033).

†Soochow University. Email: siningliu@suda.edu.cn. Address: Business School of Soochow University, Suzhou, Jiangsu, China.

‡City University of Macau. Email: yanwang@cityu.edu.mo. Address: Avenida Xian Xing Hai, No. 81–121, Centro Golden Dragon, Macau.

§Corresponding author. Peking University HSBC Business School. Email: lingxiao@phbs.pku.edu.cn. Address: University Town, Nanshan District, Shenzhen, 518055, China.

Survival of the Fittest:

A Three-Factor Model in the Currency Market

Abstract

Recent decades have witnessed a rapid growth in proposed risk factors in the international currency market. While most currency factors are evaluated using conventional asset pricing tests that can be mechanically distorted, we identify the best asset pricing model using a survival-of-the-fittest approach. Our framework treats competing currency factor models as alternative specifications of the stochastic discount factor and compares them within a Bayesian framework. We find that a parsimonious three-factor model including the dollar (DOL), carry (CAR), and business-cycle or output-gap (GAP) factors dominates the model space. This model achieves the highest marginal likelihoods, delivers positive and precisely estimated risk premia, and spans the remaining factors. Its dominance is robust across distributional assumptions, alternative factor proxies, and out-of-sample tests. Economically, the model also delivers strong pricing and investment performance, spanning most external currency portfolios and providing superior risk–return trade-offs. Overall, our results provide an economically interpretable benchmark for understanding risk compensation in foreign exchange markets.

Key Words: Currency markets; Asset pricing; Factor models; Bayesian model comparison; Heavy-tailed returns

JEL Classification: F31, G15, G12, C11

1 Introduction

Identifying which risk factors truly survive rigorous empirical scrutiny remains a central challenge in asset pricing. Since the seminal work by Lustig et al. (2011), the literature has proposed a large and growing set of candidate factors in the international currency (FX) market. While this expansion underscores the economic richness of currency risk compensation, it also complicates disciplined model evaluation. A key unresolved question is whether all proposed currency risk factors represent distinct sources of priced risk. Asset pricing theory implies the existence of a stochastic discount factor (SDF) that prices the cross-section of currency returns. Only factors that enter the SDF directly should command independent risk premia, whereas factors that are correlated with returns but spanned by the SDF are redundant. Recent evidence suggests that such redundancy may be substantial in the foreign exchange market. For example, Nucera et al. (2024) show that a small number of latent factors resembling dollar, carry, and momentum can explain a large cross-section of currency returns, while Chernov et al. (2024) find that a sizable fraction of emerging market currency volatility is unpriced. Consequently, distinguishing true sources of risk compensation from spurious or spanned factors in the FX market is essential for understanding the economic structure of currency risk premia.

To answer this question, the international asset pricing literature typically relies on standard empirical approaches, including time-series regressions, Fama–MacBeth procedures, Gibbons–Ross–Shanken (GRS) tests, and Generalized Method of Moments (GMM) methods, applied to prespecified currency portfolios formed from specific trading strategies. However, conclusions drawn from these methods are often sensitive to the choice of test assets, the ordering of tests, and distributional assumptions, leading to unstable model rankings across samples and test assets. These limitations are particularly pronounced in currency markets, where strong exchange rate comovement and heavy-tailed return behavior undermine conventional pricing tests based on Gaussian assumptions and stable second moments. Moreover, most newly proposed currency factors mainly focus on comparisons with the carry factor (Lustig et al., 2011) to justify their uniqueness in explaining currency returns, lacking assessment among the remaining currency risk factors. The proliferation of currency risk factors makes it more difficult to distinguish genuinely robust sources of risk compensation from factors whose success can be driven by sampling variation or favorable modeling assumptions.

This paper approaches factor identification as a comprehensive model comparison problem rather than a sequence of isolated hypothesis tests. Building on the Bayesian model comparison framework of Chib and Zeng (2020) and Chib et al. (2020), we evaluate economically motivated currency factor combinations as competing specifications of the stochastic discount factor (SDF). Models are treated symmetrically ex ante, and the data determine which specifications best balance pricing performance and parsimony. Our empirical implementation follows comparison protocols similar to those in Chib et al. (2024) and Qiao et al. (2022), which demonstrate the effectiveness of Bayesian model comparison in disciplining large factor spaces in U.S. and international equity markets. A key advantage of this framework is that it does not require exogenously specified test assets; instead, potentially redundant factors themselves serve as the objects to be spanned, yielding an internally consistent assessment of factor relevance.

A central feature of our analysis is the explicit modeling of heavy-tailed currency return dynamics. Ignoring tail risk is not merely a technical misspecification in international asset pricing; it obscures a fundamental source of risk compensation in the FX market associated with global downside risk and extreme macroeconomic shocks as documented in Dobrynskaya (2014), Lettau et al. (2014) and Galsband and Nitschka (2014). We therefore implement the Bayesian model scan under both Gaussian and Student-t specifications, allowing the data to discipline the role of tail risk in currency factor pricing. This approach yields more robust inference and sharper discrimination among competing models.

Applying this methodology to a broad set of widely used currency factors, we find that a parsimonious three-factor structure consisting of the dollar factor (*DOL*), the carry factor (*CAR*), and the business cycle or output gap factor (*GAP*) consistently dominates the model space. The *DOL–CAR–GAP* specification emerges as a survival-of-the-fittest model across distributional assumptions, pricing tests, and economic performance evaluations. Its relative advantage is especially pronounced under the Student-t specification, highlighting the economic importance of modeling heavy-tailed risks in currency markets. Taken together, our results provide a unified and economically interpretable benchmark for currency risk compensation and offer a disciplined foundation for future research in international asset pricing.

Our paper contributes to the literature documenting a broard set of currency risk factors, including dollar and carry factors (Lustig et al., 2011), momentum (Menkhoff et al., 2012b), value (Menkhoff et al., 2017), dollar beta (Verdelhan, 2018), FX correlation (Mueller et al., 2017), global imbalances (Corte et al., 2016), business cycles (Colacito et al., 2020), global volatility (Menkhoff et al., 2012a), downside risk (Dobrynskaya, 2014; Lettau et al., 2014; Galsband and Nitschka, 2014), liquidity risk (Mancini et al., 2013), sovereign risk (Corte et al., 2022), and FX trade volume (Cespa et al., 2022), among others. While existing studies typically assess these factors using traditional asset pricing tests, we evaluate competing factor combinations within a uniform Bayesian model comparison framework. To the best of our knowledge, this paper is the first to apply such a Bayesian approach to systematic factor evaluation in currency markets. In addition, our results provide direct supporting evidence for the role of business-cycle-related risk emphasized by Colacito et al. (2020), obtained through a novel empirical methodology.

The remainder of the paper is organized as follows. Section 2 introduces the currency pricing framework and the Bayesian model comparison methodology. Section 3 describes the data and presents the empirical results. Section 4 evaluates the economic performance of the best specification, the *DOL–CAR–GAP* model. Section 5 concludes.

2 Methodology

This section describes the Bayesian model scan framework used to identify the set of pricing factors entering the stochastic discount factor (SDF). Building on Chib and Zeng (2020) and Chib et al. (2020), we conduct a comprehensive comparison over all factor combinations. Each candidate specification corresponds to a distinct stochastic discount factor, and models are evaluated using posterior model probabilities and marginal likelihoods under alternative distributional assumptions. Treating factor identification as a model comparison problem provides a unified and internally consistent criterion for ranking competing pricing kernels

2.1 Model Specification

Formally, consider a universe of N candidate currency factors. We specify a candidate model $\mathcal{M}_s$, indexed by $s = 1, \ldots, S$, as a unique partition of these factors into two disjoint vectors:

1. The vector of pricing factors: $\boldsymbol{z}_t \in \mathbb{R}^{n_z}$, and
2. The vector of remaining factors: $\boldsymbol{q}_t \in \mathbb{R}^{n_q}$,

satisfying the condition $N = n_z + n_q$. Given the theoretical requirement that a valid SDF must contain at least one pricing factor (i.e., $n_z \geq 1$), the total number of competing models is $S = 2^N - 1$.

Let $\boldsymbol{f}_t = (\boldsymbol{z}_t', \boldsymbol{q}_t')'$ denote the joint factor vector. The central tenet of our analysis is that the no-arbitrage condition imposes a strict structural restriction on the data generating process of $\boldsymbol{f}_t$. Specifically, under the hypothesis that $\boldsymbol{z}_t$ constitutes the true pricing kernel, the reduced-form model for specification $\mathcal{M}_s$ is governed by the following system of equations:

$$\boldsymbol{z}_t = \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t^{(z)} \tag{1}$$

$$\boldsymbol{q}_t = \mathbf{B}\boldsymbol{z}_t + \boldsymbol{\varepsilon}_t^{(q|z)} \tag{2}$$

where $\boldsymbol{\delta}$ represents the vector of risk premia for the pricing factors.

The defining feature of this system is the absence of an intercept vector (or $\boldsymbol{\alpha}$) in Equation (2). This zero-intercept restriction is not an arbitrary statistical assumption; it is a direct derivation from the fundamental pricing condition $\mathbb{E}[M_t \boldsymbol{f}_t] = 0$. Because the redundant factors $\boldsymbol{q}_t$ do not enter the SDF independently, their expected excess returns must be fully spanned by the pricing factors $\boldsymbol{z}_t$. Consequently, any non-zero intercept in Equation (2) would imply an arbitrage opportunity, contradicting the premise that $\boldsymbol{z}_t$ serves as the sole pricing kernel.

To implement this framework empirically, we model the joint distribution of the candidate factors $\boldsymbol{f}_t$ under two distinct distributional assumptions to ensure robustness.

(i) Gaussian Framework Under the assumption of multivariate normality, the errors follow a block-diagonal normal distribution:

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t^{(z)} \\ \boldsymbol{\varepsilon}_t^{(q|z)} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \boldsymbol{\Sigma}_z & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{q|z} \end{pmatrix}\right) \tag{3}$$

Here, $\boldsymbol{\Sigma}_z$ is the covariance matrix of pricing factors while $\boldsymbol{\Sigma}_{q|z} \in \mathbb{R}^{n_q \times n_q}$ is the conditional covariance matrix of nonpricing factors. The block-diagonal structure implies that, conditional on the pricing factors $\boldsymbol{z}_t$, the residuals of the nonpricing factors contain no further pricing information.

(ii) Student-t Framework To allow for fat tails and greater robustness to outliers, we also consider the case where the joint distribution of the currency factor vector follows a multivariate

Student-t distribution:

$$\boldsymbol{f}_t \sim \mathcal{S}t(\boldsymbol{\mu}, \mathbf{V}, \nu), \tag{4}$$

where $\boldsymbol{\mu} \in \mathbb{R}^N$ is the mean vector and $\mathbf{V} \in \mathbb{R}^{N \times N}$ is a positive definite scale matrix. The degrees of freedom parameter ν is fixed. Using the scale mixture representation of the Student-t distribution, we rewrite the model hierarchically. We introduce a latent scaling variable ω_t such that:

$$\boldsymbol{f}_t \mid \omega_t \sim \mathcal{N}(\boldsymbol{\mu}, \omega_t^{-1}\mathbf{V}) \tag{5}$$

$$\omega_t \sim \text{Gamma}(\nu/2, \nu/2) \tag{6}$$

In this setup, the linear pricing restrictions defined in the previous subsection remain identical, but the error residuals are now conditionally Gaussian given ω_t:

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t^{(z)} \\ \boldsymbol{\varepsilon}_t^{(q|z)} \end{pmatrix} \Bigg| \omega_t \sim \mathcal{N}\left(\mathbf{0}, \omega_t^{-1}\begin{pmatrix} \boldsymbol{\Sigma}_z & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{q|z} \end{pmatrix}\right) \tag{7}$$

where the scalar ω_t captures time-varying volatility and outlier events. The conditional covariance $\boldsymbol{\Sigma}_{q|z}$ relates to the scale matrix $\mathbf{V}$ via the identity $\boldsymbol{\Sigma}_{q|z} = \mathbf{V}_{qq} - \mathbf{V}_{qz}\mathbf{V}_{zz}^{-1}\mathbf{V}_{qz}'$.

Our objective is to estimate the parameters of both the Gaussian and Student-t specifications and compare them using the Bayesian approach.

2.2 Prior Specification

To ensure an unbiased "survival of the fittest" contest, the specification of prior distributions requires careful theoretical grounding. Following the insight of Chib et al. (2020), we first observe that the parameters across distinct candidate models are not independent entities; rather, they are intrinsically related by change-of-variable formulas. Consequently, the priors for competing specifications should not be defined in isolation, but effectively derived from a single, common underlying prior. This theoretical consistency is paramount to ensure that the final comparison results are driven by the data's support for the pricing restrictions, rather than by arbitrary differences in prior density heights.

Furthermore, in the context of marginal likelihood comparison, proper and informative priors are strictly superior to arbitrary non-informative (improper) priors. The use of diffuse or improper priors leads to "Bartlett's Paradox," where the marginal likelihood becomes sensitive to undefined normalization constants, rendering Bayes factors indeterminate.

To satisfy these requirements, we adopt the Training Sample Prior (TSP) approach. We partition the full dataset $\mathcal{D}$ into two disjoint sets: a training sample $\mathcal{D}_0$ (comprising the first T_0 observations) and an estimation sample $\mathcal{D}_{est}$ (comprising the remaining $T - T_0$ observations). By anchoring the priors in the same initial data subset $\mathcal{D}_0$ for all models, this method naturally enforces the necessary change-of-variable consistency. Analytically, this implies that the marginal likelihood is interpreted as the *predictive density* of the estimation sample, conditional on the information in the training sample.

2.3 Model Comparison

The final ranking of models is determined by the marginal likelihood, $m(\mathcal{D}_{est} \mid \mathcal{M}_s)$, $s = 1, 2, \ldots, S$, which represents the probability of observing the estimation data under model $\mathcal{M}_s$, integrated over all parameter uncertainty. We assume a uniform prior over the model space, $\Pr(\mathcal{M}_s) = 1/S$, ensuring the ranking is driven solely by data fit.

Gaussian Framework (Analytical Solution) A major computational advantage of the Gaussian framework under the TSP is the existence of a closed-form solution. The log-marginal likelihood for a specific model $\mathcal{M}_s$ can be decomposed into two intuitive components:

$$
\begin{aligned}
&\ln m_s(\mathcal{D}_{est} \mid \mathcal{M}_s) \\
&= \underbrace{-\frac{Tn_{z,j}}{2}\ln\pi - \frac{n_{z,j}}{2}\ln(T\kappa_s+1) - \frac{T+n_{z,j}-N}{2}\ln|\Psi_s| + \ln\Gamma_{n_{z,j}}\left(\frac{T+n_{z,j}-N}{2}\right)}_{\text{Pricing Kernel Fit}} \\
&\quad \underbrace{-\frac{(N-n_{z,j})(T-n_{z,j})}{2}\ln\pi - \frac{N-n_{z,j}}{2}\ln|W_s^*| - \frac{T}{2}\ln|\Psi_s^*| + \ln\Gamma_{N-L_s}\left(\frac{T}{2}\right)}_{\text{Spanning Efficiency of Remaining Factors}}
\end{aligned} \tag{8}
$$

where κ_s is chosen by the prior, Ψ_s measures the sum of squared errors for the pricing factors, and Ψ_s^* measures the residual variance of the nonpricing factors.

- The **first component** rewards the model for explaining the dynamics of the chosen pricing factors $\boldsymbol{z}_t$ parsimoniously.
- The **second component** penalizes the model if the nonpricing factors $\boldsymbol{q}_t$ are not well-spanned by $\boldsymbol{z}_t$. If the "redundant" factors have large alphas (large Ψ_s^*), this term drops, lowering the model's rank.

Student-t Framework (MCMC Algorithm) In the Student-t framework, the presence of the latent volatility variable ω_t precludes a closed-form solution. Instead, we estimate the log-marginal likelihood using the method developed by Chib (1995). This approach rearranges Bayes' theorem to evaluate the marginal likelihood at a specific high-density point $\boldsymbol{\theta}^*$ (typically the posterior mean):

$$
\ln p(\mathcal{D}_{est} \mid \mathcal{M}_s) = \underbrace{\ln p(\mathcal{D}_{est} \mid \boldsymbol{\theta}^*, \mathcal{M}_s)}_{\text{Log Likelihood}} + \underbrace{\ln \pi(\boldsymbol{\theta}^* \mid \mathcal{M}_s)}_{\text{Log Prior}} - \underbrace{\ln p(\boldsymbol{\theta}^* \mid \mathcal{D}_{est}, \mathcal{M}_s)}_{\text{Log Posterior}} \tag{9}
$$

The likelihood and prior ordinates are calculated analytically, while the posterior ordinate $\ln p(\boldsymbol{\theta}^* \mid \mathcal{D}_{est}, \mathcal{D}_0, \mathcal{M}_s)$ is estimated via the output of the MCMC simulation using the reduced-run method. This rigorous computation ensures that our selection of the "fittest" model is robust to the distributional realities of currency markets, such as heavy tails and volatility clustering.

We conclude the methodology by summarizing the selection criterion. For each candidate model $\mathcal{M}_s$, we compute the marginal likelihood $m(\mathcal{D}_{\text{est}} \mid \mathcal{M}_s)$ under both the Gaussian and Student-t specifications using the training-sample prior. Given the uniform prior over the model space, $\Pr(\mathcal{M}_s) = 1/S$, the posterior model probabilities are proportional to these marginal likelihoods, so that ranking models by posterior probability is equivalent to ranking them by log marginal likelihood. This Bayesian model scan therefore delivers a definitive, data-driven hierarchy of competing specifications in which the "fittest" model is the one that maximizes the marginal likelihood, balancing goodness-of-fit, parsimony, and spanning efficiency. In the next section, we apply this framework to the currency factor zoo and determine which subset of factors is most strongly supported by the data under this Bayesian model comparison approach.

3 Empirical Results

3.1 Candidate Currency Factors

In our empirical analysis, we consider nine candidate currency factors: *DOL*, *DDOL*, *CAR*, *MOM*, *DB*, *VAL*, *FXC*, *GAP*, and *IMB*, with detailed description, construction and reference provided in Table 1. These factors are widely used in the international finance literature as key sources of risk compensation and are hypothesized to capture distinct yet complementary economic mechanisms underlying currency excess returns. Taken together, they are designed to span global risk sentiment, monetary policy divergence, valuation adjustments, macroeconomic momentum, and systemic funding vulnerabilities.

The **Dollar factor (DOL)** and its conditional counterpart, the **Dollar Carry factor (DDOL)**, serve as the baseline global market factors in our analysis. **DOL** measures the global price of shorting U.S. dollar while longing all foreign currencies, reflecting the broad risk sentiment between them (Lustig et al., 2011). **DDOL** refines this strategy by conditioning on the sign of the global interest-rate differential (Lustig et al., 2014), highlighting the cyclical nature of **DOL** by switching between long and short positions depending on funding conditions. Given that *DDOL* is a conditional transformation of the aggregate market factor, we treat *DOL* and *DDOL* as alternative benchmarks and preclude their simultaneous inclusion in any model specification.

Distinct from market-average risks, the **Carry factor (CAR)** compensates investors for providing capital to high–interest-rate currencies that become particularly risky in global downturns, consistent with violations of uncovered interest parity (Lustig et al., 2011). The **Momentum factor (MOM)** reflects investor under-reaction and trend-chasing behavior: currencies with strong past performance continue to appreciate while previous losers underperform, generating a persistent return spread (Menkhoff et al., 2012b). The **Dollar Beta factor (DB)** captures heterogeneous exposure to systematic dollar risk; currencies with high sensitivity to **DOL** behave analogously to high-beta equities and therefore command higher expected returns when global risk aversion rises (Verdelhan, 2018). The **Value factor (VAL)** reflects currency misalignment relative to purchasing power parity, as undervalued currencies tend to appreciate when economic fundamentals revert toward long-run equilibrium (Menkhoff et al., 2017).

The remaining macro-structural factors link currency returns to broader sources of systemic risk. The **FX Correlation factor (FXC)** prices variation in co-movement across currencies, which increases sharply during crises and erodes diversification benefits when they are most valuable (Mueller et al., 2017). The business cycle or **Output Gap factor (GAP)** relates exchange rate dynamics to business-cycle fluctuations, rewarding currencies issued by economies operating above potential output and penalizing those tied to recessionary conditions (Colacito et al., 2020). Finally, the **Global Imbalance factor (IMB)** differentiates safe-haven currencies from risky debtor-country currencies and compensates investors for exposure to external funding stress and current-account fragility (Corte et al., 2016).

Currency excess returns are calculated using future spot rates and current forward rates as follows:

$$r_{t+1} = f_t - s_{t+1}, \tag{10}$$

where r_{t+1} denotes the currency excess return earned by investors who short the U.S. dollar and go long foreign currency in month t; s_{t+1} is the logarithm of the spot exchange rate in month $t+1$; and f_t is the logarithm of the one-month forward rate in month t. Spot and forward rates are obtained from Reuters and Barclays via Datastream. We exclude turmoil episodes during which the data are deemed unreliable, following Lustig et al. (2011). Both spot and forward rates are defined as the amount of foreign currency per unit of U.S. dollar. Thus, an increase in the exchange rate indicates an appreciation of the U.S. dollar and a depreciation of the foreign currency.

The countries or regions included in our sample are Australia, Austria, Belgium, Brazil, Canada, Croatia, Cyprus, the Czech Republic, Denmark, the euro area, Finland, France, Germany, Greece, Hong Kong SAR, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Kuwait, Malaysia, Mexico, the Netherlands, New Zealand, Norway, the Philippines, Poland, Portugal, Russia, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Thailand, and the United Kingdom. Upon joining the euro area, countries are automatically removed from the sample, with only the euro retained.

Our empirical analysis draws on a comprehensive monthly dataset spanning January 1991 through March 2024. Table 2 reports the descriptive statistics for currency returns of the candidate factors above, including their annualized means, standard deviations, skewness, kurtosis, maximum and minimum values in percentage, which are comparable to the empirical evidence documented in the literature. Notably, across all factors, the kurtosis values substantially exceed three, indicating pronounced tail thickness. This is a well-documented feature of currency return series, a finding illustrated, for example, by the carry portfolios in Dobrynskaya (2014). This systematic departure from Gaussian behavior provides a clear rationale for employing a heavy-tailed assumption to capture extreme movements more accurately.

For the purposes of out-of-sample evaluation, the final twelve months of data (April 2023 to March 2024) are reserved for predictive ability analysis. From the remaining observations, the first 15% constitute the training sample used to construct the prior distributions, while the subsequent observations form the estimation sample for posterior inference and model comparison.

3.2 Economically Motivated Model Scan

Having established the full set of nine candidate currency factors, we now describe the economically motivated model scan underpinning our Bayesian model comparison. Because *DOL* and *DDOL* represent alternative proxies for the global market factor, we treat them as substitutes. To maintain a consistent benchmark, our primary analysis focuses on the candidate set containing *DOL* and the seven characteristic-based factors ($N = 8$), reserving *DDOL* for separate robustness checks. This design ensures that the model space spans a wide range of theoretically grounded currency risk exposures while avoiding redundancy across overlapping market proxies.

Although the Bayesian framework permits a fully agnostic search over all nonempty subsets of these factors, our main analysis imposes an economically motivated structure consistent with prevailing empirical practice. As emphasized by Corte et al. (2016), leading studies routinely employ a two-factor stochastic discount factor (SDF) in which an average market component, proxied by the average excess return on a long-foreign–short-dollar position (*DOL*), plays a central and indispensable role. This currency market risk factor is viewed as the primary driver of common variation in currency excess returns and therefore forms the backbone of most empirical pricing models.

In line with this literature, we require *DOL* to be included in every candidate specification.[1]

[1] Related treatments in the equity market literature also maintain the market factor in all candidate specifications,

The remaining seven factors are freely selectable, allowing the Bayesian procedure to determine whether they contribute incremental explanatory power beyond the market component. This restriction yields a feasible space of $S = 2^7 = 128$ distinct model specifications, in contrast to a fully unrestricted search of the candidate set that would involve $2^8 - 1$ combinations.

Each model is estimated under both Gaussian and Student-t assumptions, with degrees of freedom fixed at $\nu = 5$.[2] This economically motivated yet comprehensive model space provides a disciplined platform for identifying whether a specific model consistently dominates across varying distributional frameworks.

The Best Model: *DOL–CAR–GAP* To identify the most informative combination of currency risk factors, we conduct a comprehensive Bayesian model comparison over the restricted model space that enforces inclusion of the dollar factor (*DOL*) in every specification. Each model among the 128 feasible models is evaluated using its marginal likelihood, which quantifies its ability to explain the observed cross-section of currency excess returns after integrating out uncertainty in model parameters. By comparing marginal likelihoods across all candidates, the Bayesian framework provides a principled mechanism for balancing model fit and complexity, thereby selecting the specification with the highest degree of support from the data.

As noted in the methodology section, the prior probability assigned to each model in the restricted model space is uniform, given by $\mathbb{P}(\mathcal{M}_s) = 1/128 \approx 0.78\%$. After computing the (log) marginal likelihoods for all candidate models, the posterior model probabilities are obtained as

$$\mathbb{P}(\mathcal{M}_s \mid \mathcal{D}_{\text{est}}) = \frac{m_s(\mathcal{D}_{\text{est}} \mid \mathcal{M}_s)}{\sum_{s=1}^{128} m_s(\mathcal{D}_{\text{est}} \mid \mathcal{M}_s)}, \qquad s = 1, 2, \ldots, 128, \tag{11}$$

where $m_s(\mathcal{D}_{\text{est}} \mid \mathcal{M}_s)$ denotes the marginal likelihood of the estimation sample $\mathcal{D}_{\text{est}}$ under

as discussed in Barillas and Shanken (2018) and Chib et al. (2020).

[2]We adopt $\nu = 5$ as our baseline for the Student-t specification, providing a flexible yet empirically stable representation of heavy tails in currency returns. Robustness checks with alternative values of ν yield qualitatively similar results.

model $\mathcal{M}_s$. Figure 1 displays the resulting posterior probabilities for all ranked models under the *DOL*-restricted scan across both the Gaussian and Student-t specifications, providing a transparent view of how the data update the uniform prior beliefs over the economically prior–driven restricted model space.

In both distributional settings, the posterior distribution is sharply concentrated on a small set of leading models, providing strong Bayesian evidence against model dispersion and indicating that only a limited number of factor combinations receive economically meaningful probability mass. A single specification, the three-factor *DOL–CAR–GAP* model, dominates the posterior distribution under both distributional specifications. Under the Gaussian specification in subplot (a), the posterior probability of the *DOL–CAR–GAP* model exceeds 15%, nearly twenty times its uniform prior benchmark. Under the Student-t specification in subplot (b), the concentration is even stronger, with posterior probability exceeding 25%. This dramatic amplification from prior to posterior delivers compelling evidence for the stability and dominance of the *DOL–CAR–GAP* model as the leading explanation for cross-sectional variation in currency excess returns.

Beyond the best model, the remaining top-ranked alternatives include minor extensions such as augmenting the specification with *DB* or *VAL*. However, their posterior weights decline sharply, and no competing model comes close to the probability mass assigned to the *DOL–CAR–GAP* specification under either distributional assumption. The stability of this pattern across both Gaussian and Student-t specifications indicates that the prominence of the *DOL–CAR–GAP* model is not an artifact of distributional assumptions or sampling variability, but rather reflects genuine explanatory power in capturing the cross-sectional variation in currency excess returns.

Taken together, these results establish the *DOL–CAR–GAP* specification as the uniquely dominant factor model within the economically motivated model scan.

Comparison with competing models To gauge whether the Bayesian best model reflects meaningful economic structure rather than a statistical artifact, we benchmark it against prominent currency factor specifications proposed in the literature. These include the *DOL–CAR* model of Lustig et al. (2011), the *DOL–MOM* and *DOL–CAR–MOM* models of Menkhoff et al. (2012b) and Nucera et al. (2024), and the *DOL–IMB* model of Corte et al. (2016), among others. Most of these specifications are nested within our *DOL*-restricted model space, while a small number of alternatives that replace *DOL* with another factor (such as *DB*) lie in the corresponding unrestricted space. In all cases, the models are evaluated under the same prior structure, allowing us to compute marginal likelihoods on a comparable basis.

Table 3 reports the resulting log marginal likelihood values for the Bayesian best model and the competitors. The *DOL–CAR–GAP* model attains the highest log marginal likelihood of 6408.038 under the Gaussian specification and 6560.042 under the Student-t specification, confirming its ranking as the Bayesian best model within the economically prior–driven candidate set. Relative to the competing models, *DOL–CAR–GAP* achieves uniformly higher log marginal likelihoods: it dominates the classical *DOL–CAR* benchmark, the *DOL–MOM* and *DOL–CAR–MOM* models, and the *DOL–IMB* model, as well as alternative structures that substitute *DOL* with *DB*. No competing model approaches the posterior probability or marginal likelihood of *DOL–CAR–GAP* under either distributional assumption.

Evidence for Heavy Tails The fact that the *DOL–CAR–GAP* specification remains the best-performing model under both Gaussian and Student-t likelihoods, and attains a noticeably higher log marginal likelihood under the Student-t specification, provides direct evidence that heavy-tailed behavior is an important feature of currency factor dynamics. In other words, allowing for fat tails in the return distribution does not change the identity of the preferred model, but it strengthens the degree of Bayesian support in its favor. This pattern is consistent with the well-documented empirical properties of currency returns and evidence in Table 2, which typically exhibit excess kurtosis and pronounced tail risk. By accommodating these features through a Student-t specification, the model better captures extreme movements in factor returns, particularly those associated with episodes of funding stress or sudden shifts in global risk sentiment. Overall, the comparison across marginal likelihoods suggests that heavy-tailed distributions offer a more realistic description of currency factor returns and that the dominance of the *DOL–CAR–GAP* specification is, if anything, understated under the Gaussian benchmark.

Choice of ν We also examine the robustness of our benchmark Student-t specification with respect to the choice of the degrees of freedom parameter ν. Our baseline model scan adopts a Student-t likelihood with $\nu = 5$, a value commonly used in the empirical asset pricing literature.

To evaluate the sensitivity of our findings, Table 4 reports the log-marginal likelihood of the *DOL–CAR–GAP* model under Student-t distributions with degrees of freedom ranging from 2.1 to ∞ (the Gaussian limit). As ν increases, the log-marginal likelihood first rises and then declines, displaying an inverted U-shaped pattern. The highest values occur around $\nu = 4.5$ and $\nu = 5$, which supports our baseline choice of $\nu = 5$ in the benchmark scan. At the same time, the log-marginal likelihood does not fluctuate dramatically across a wide range of ν: the *DOL–CAR–GAP* model continues to deliver log-marginal likelihoods above 6500 for ν between 3 and 10.

Overall, these results indicate that our main conclusions are not sensitive to a finely tuned choice of the tail parameter. The *DOL–CAR–GAP* specification remains the best-performing model over a broad and empirically plausible range of degrees of freedom, and the evidence in favor of a heavy-tailed Student-t specification is robust.

Economic interpretation Economically, the *DOL–CAR–GAP* model incorporates three distinct sources of priced currency risk. *DOL*, the dollar factor, captures broad risk sentiment and flight-to-safety dynamics. *CAR* the carry factor, which is proxied by the excess return of longing currencies with high interest rates while shorting those with low interest rates, compensates investors for exposure to volatility risk in bad times (Lustig et al., 2011). *GAP*, constructed by longing currencies with high output gaps and shorting those with low gaps, captures global business-cycle divergence and macroeconomic momentum (Colacito et al., 2020). Thus, although selected through a fully data–driven Bayesian search, the best–performing specification corresponds to a theoretically coherent structure directly aligned with recent advances in currency asset pricing research. Furthermore, our results underscore the importance of modeling the distributional tails: the marginal likelihoods decisively favor the Student-t specification, a finding consistent with the heavy-tailed data patterns and frequent extreme shocks observed in global currency markets.

3.3 Posterior Risk Premia Distributions for the Selected Factors

Having identified *DOL–CAR–GAP* as the Bayesian best specification, we next examine the posterior distributions of the associated risk premia to verify whether these factors command statistically significant compensation for risk. While this objective parallels the standard Fama and MacBeth (1973) procedure, our Bayesian framework offers the distinct advantage of flexible distributional modeling. Throughout this analysis, we work under the Student-t likelihood with $\nu = 5$, which is favored by the model comparison results presented earlier. The higher marginal likelihoods obtained under the Student-t specification indicate that heavy-tailed innovations provide a more realistic description of currency factor returns; accordingly, all posterior inference on the risk premia is reported under this preferred heavy–tailed setting.

Figure 2 displays the posterior distributions of the monthly risk premia for the three currency factors in the *DOL–CAR–GAP* model. Each panel shows the posterior density of a given component of $\boldsymbol{\delta}$ together with its posterior mean (vertical red dashed line). All three distributions are tightly centered on positive values and exhibit relatively limited dispersion, implying that the bulk of the posterior mass lies well away from zero. In particular, the posterior credible intervals for the risk premia of *DOL*, and specifically *CAR* and *GAP*, exclude zero with high probability, providing strong evidence that each factor, especially *CAR* and *GAP*, commands a statistically and economically meaningful price of risk in the FX market.

From an inferential perspective, examining these posterior risk premia plays a role analogous to the second-stage regressions in the Fama and MacBeth procedure. In a frequentist setting, the significance of the price-of-risk parameters $\boldsymbol{\delta}$ is assessed via t-statistics and their associated p-values. In our Bayesian framework, significance is instead evaluated through the shape of the posterior distributions, specifically by whether the posterior mass for each component of $\boldsymbol{\delta}$ is concentrated away from zero and by the extent to which credible intervals exclude zero. Thus, Figure 2 provides a fully probabilistic assessment of factor pricing, confirming that the three pillars of the selected model, the dollar, carry risk, and business-cycle gap risk, are each robustly compensated in equilibrium.

3.4 Internal Spanning Test of the Losing Factors

Our Bayesian, economically prior–driven model scan naturally partitions the candidate currency factors into two groups: the "winning" factors that constitute the optimal pricing kernel {DOL, CAR, GAP}, and the complementary set of "losing" (remaining or redundant) factors given by {MOM, DB, VAL, FXC, IMB}. Within our framework, these losing factors serve as natural internal test assets. If the *DOL–CAR–GAP* specification effectively spans the currency pricing kernel, the losing factors should not earn systematic residual returns (alphas) once projected onto the three winners. We test this implication using both a direct Bayesian model comparison and the classical frequentist GRS test.

Bayesian spanning test For each excluded factor $q_{j,t}$, we compare two competing regressions on the winning factors. The unrestricted model allows for a non–zero intercept (alpha),

$$\mathcal{M}_1: \quad q_{j,t} = \alpha_j + \beta_{j,\mathrm{DOL}}\,\mathrm{DOL}_t + \beta_{j,\mathrm{CAR}}\,\mathrm{CAR}_t + \beta_{j,\mathrm{GAP}}\,\mathrm{GAP}_t + \varepsilon_{j,t}, \tag{12}$$

while the restricted model imposes the pricing restriction $\alpha_j = 0$,

$$\mathcal{M}_0: \quad q_{j,t} = \beta_{j,\mathrm{DOL}}\,\mathrm{DOL}_t + \beta_{j,\mathrm{CAR}}\,\mathrm{CAR}_t + \beta_{j,\mathrm{GAP}}\,\mathrm{GAP}_t + \varepsilon_{j,t}. \tag{13}$$

In line with our results in Section 3.2, we model the error terms with a Student-t distribution with $\nu = 5$ degrees of freedom to capture the heavy tails in candidate factor returns. Let m_0 and m_1 denote the marginal likelihoods of $\mathcal{M}_0$ and $\mathcal{M}_1$ for a given losing factor. The Bayes factor in favor of the pricing restriction is

$$\mathrm{BF}_{01} = \frac{m_0}{m_1}. \tag{14}$$

Values $\mathrm{BF}_{01} > 1$ indicate that the data favor the zero–alpha restriction; equivalently, $\ln(\mathrm{BF}_{01}) > 0$ implies support for spanning.

Table 5 reports the log marginal likelihoods and the implied Bayes factors for all excluded factors. For each of *MOM*, *DB*, *VAL*, *FXC*, and *IMB*, the restricted model $\mathcal{M}_0$ attains a higher marginal likelihood than the unrestricted model, so that $\mathrm{BF}_{01} > 1$ as well as $\ln(\mathrm{BF}_{01}) > 0$ in every case. The evidence is strongest for the Momentum factor (*MOM*), with a Bayes factor of about 6.99, and more moderate but still larger than 1 for the remaining factors (Bayes factors between roughly 1.36 and 2.31). Taken together, these results indicate that the data systematically prefer the zero–alpha specification and that the losing factors are well spanned by the *DOL–CAR–GAP* kernel.

GRS test To complement the Bayesian analysis with a standard frequentist benchmark, we implement the multivariate spanning test of Gibbons et al. (1989) (GRS). For each excluded factor $q_{j,t} \in \{\text{MOM}, \text{DB}, \text{VAL}, \text{FXC}, \text{IMB}\}$, we estimate

$$q_{j,t} = \alpha_j + \beta_{j,\text{DOL}}\,\text{DOL}_t + \beta_{j,\text{CAR}}\,\text{CAR}_t + \beta_{j,\text{GAP}}\,\text{GAP}_t + u_{j,t}, \tag{15}$$

where $u_{j,t} \sim \mathcal{N}(0, \sigma_j^2)$ represents the Gaussian pricing error. We test the joint null hypothesis that all intercepts are zero:

$$H_0 : \alpha_{\text{MOM}} = \alpha_{\text{DB}} = \alpha_{\text{VAL}} = \alpha_{\text{FXC}} = \alpha_{\text{IMB}} = 0. \tag{16}$$

In our in–sample data, the GRS statistic equals 1.79 with degrees of freedom $(\text{df}_1, \text{df}_2) = (5, 379)$, yielding a p–value of 0.114. We therefore fail to reject the joint zero–alpha null at conventional significance levels (5% or 10%). From a frequentist perspective, there is no evidence that the losing factors earn abnormal returns once their exposures to *DOL*, *CAR*, and *GAP* are taken into account.

Synthesis of spanning evidence. The Bayesian model comparison and the GRS spanning test lead to a coherent conclusion. The *DOL–CAR–GAP* specification not only maximizes the marginal likelihood over the currency factor zoo, but also successfully prices the complementary factors in the sense that their residual alphas are jointly indistinguishable from zero. This internal consistency strengthens the interpretation of the losing factors as genuinely redundant and reinforces the view that *DOL–CAR–GAP* provides an efficient and parsimonious representation of the currency pricing kernel.

3.5 Out-of-sample (OOS) Predictive Performance

Following Chib et al. (2024), we evaluate the out-of-sample (OOS) predictive performance of alternative factor specifications using the predictive likelihood, which is the standard Bayesian measure of forecast performance.[3]

Table 6 reports the OOS log predictive likelihoods for the in-sample Bayesian best model and the main competing currency factor specifications using the reserved final 12 months (April 2023 to March 2024) observations. Consistent with the in-sample evidence in Table 3, the three-factor *DOL–CAR–GAP* model achieves the highest predictive likelihood under both the Gaussian and Student-t assumptions. Taken together, the in-sample and out-of-sample results demonstrate that the *DOL–CAR–GAP* specification selected by the Bayesian model scan not only provides the best fit to the historical data, but also delivers superior predictive performance relative to other leading currency factor models.

3.6 Robustness Check: Replacing *DOL* with *DDOL* in the Model Scan

As a robustness exercise, we replicate the economically motivated model scan after replacing the dollar factor *DOL* with its conditional counterpart *DDOL*. Figure 3 reports the resulting posterior model probabilities under both the Gaussian and Student-t likelihoods. Across specifications, the highest posterior probability model is consistently *DDOL–CAR–GAP*. This outcome closely mirrors the original model scan results based on *DOL*, indicating that our main findings are not sensitive to the particular choice of the dollar factor.

More broadly, the stability of the selected three-factor model confirms that the explanatory role of the Carry and Output Gap factors is intrinsic to the data rather than an artifact of how the global currency component is measured.

[3] For a given realization of future data, the predictive likelihood, like the marginal likelihood, evaluates to a scalar and thus provides a natural basis for ranking models across the candidate space.

Summary of Empirical Results

In this section, we introduced a comprehensive set of nine candidate currency factors. We documented that their return distributions display strong excess kurtosis, which motivates the use of a Student-t specification in our Bayesian framework. Within an economically motivated model space that always includes the dollar factor *DOL*, the Bayesian model scan selects *DOL–CAR–GAP* as the clear best model under both Gaussian and Student-t assumptions. Its posterior probability far exceeds its uniform prior weight, and its marginal likelihood is higher than those of leading benchmark models such as *DOL–CAR*, *DOL–MOM*, *DOL–CAR–MOM*, and *DOL–IMB*. The result is robust to alternative tail parameters and to replacing *DOL* with its conditional analogue *DDOL*.

We further show that all three factors in the *DOL–CAR–GAP* model, especially the carry and output gap factors, command positive and precisely estimated risk premia under the Student-t specification. In contrast, the remaining "losing" factors are well explained by this pricing kernel: Bayesian spanning tests indicate no intercepts, while the GRS test fails to reject the joint null hypothesis that pricing errors are zero. Out-of-sample predictive likelihoods further confirm that *DOL–CAR–GAP* delivers the strongest forecasting performance among competing models. Taken together, these findings establish *DOL–CAR–GAP* as a parsimonious, statistically strong, and economically meaningful benchmark.

4 Economic Performance of the *DOL–CAR–GAP* Model

In the previous section, our Bayesian model comparison identified the specification comprising the dollar (*DOL*), carry (*CAR*), and output gap (*GAP*) factors as the statistically dominant candidate. The Student-t specification with fat tails delivers higher marginal likelihoods than the Gaussian benchmark, and the main findings are robust to alternative degrees of freedom as well as to model scans based on *DDOL* or without inclusion restrictions. We also showed that the winning three-factor model successfully spans the remaining currency factors from both Bayesian and frequentist perspectives, and that it achieves superior out-of-sample predictive performance as measured by the predictive likelihood.

However, statistical dominance alone does not guarantee economic significance or structural stability. In this section, we therefore conduct a comprehensive economic evaluation of the *DOL–CAR–GAP* model relative to the leading currency factor models in the literature.[4] We first examine its pricing performance using Bayesian spanning tests and regression-based diagnostics, respectively. We then evaluate its investment performance both in-sample by comparing Minimum-Variance Efficient Frontiers across different model specifications and out-of-sample by assessing predictive consistency through out-of-sample Sharpe ratios.

4.1 Pricing Performance

As noted in the introduction, a distinguishing feature of our Bayesian model scan is that the selection of the superior specification is independent of the specific test assets employed for evaluation. Although the *DOL–CAR–GAP* model was selected based on marginal likelihoods of the factors themselves, we must still rigorously validate its ability to price external currency portfolios.

We employ a comprehensive universe of $K = 34$ test portfolios to challenge the model. These assets represent a diverse set of currency strategies, ensuring that our results are not driven by a specific anomaly. The test set comprises: 5 Carry, 5 Value, 5 Momentum, 4 FX Correlation, 6 Dollar Beta, 4 Imbalance, and 5 Output Gap portfolios. Their constructions are described below:

Carry portfolios (Lustig et al., 2011): We sort currencies into five portfolios based on forward discounts (interest rate differentials).

[4] For completeness, we also report parallel results for the *DDOL*-based analogue of the model.

Value portfolios (Menkhoff et al., 2017): We sort currencies into five portfolios based on 5-year changes in real exchange rates.

Momentum portfolios (Menkhoff et al., 2012b): We sort currencies into five portfolios based on past 1-month currency returns.

FX correlation portfolios (Mueller et al., 2017): We sort currencies into four portfolios based on the loadings of currency returns on innovations in the FX correlation dispersion measure.

Dollar Beta portfolios (Verdelhan, 2018): We sort currencies into six portfolios based on loadings on the *DOL* factor. Each portfolio takes a long (short) position if the median forward discount rate of developed currencies is positive (negative).

Imbalance portfolios (Corte et al., 2016): We first sort currencies into two portfolios based on the net foreign asset to GDP ratio, and subsequently sort each portfolio into two baskets using the share of foreign liabilities in domestic currency.

Output Gap portfolios (Colacito et al., 2020): We sort currencies into five portfolios based on output gaps.

Our full sample dataset spans the period from January 1991 to March 2024, yielding a monthly panel that is fully aligned with the candidate currency factors introduced in Section 3.1. For each month in this sample, we observe returns of both the nine candidate currency factors (including the two market-proxy series *DOL* and *DDOL*) and the 34 test portfolios described above, all constructed using the same underlying currency universe and data filters. Throughout this section, all pricing tests for the benchmark *DOL–CAR–GAP* model, its analogue *DDOL–CAR–GAP*, and the competing specifications are conducted on this full sample of monthly excess returns, so that comparisons are made on a common information set and are directly comparable across models.

4.1.1 Bayesian Spanning Test

To rigorously assess the pricing performance of the *DOL–CAR–GAP* model and its closely related alternative *DDOL–CAR–GAP*, we employ a Bayesian model comparison framework. Consistent with our analysis in Section 3.2, we consider both a Gaussian specification and a Student-t specification for the error terms, fixing the degrees of freedom at $\nu = 5$ to accommodate the heavy tails observed in currency returns.

Take the *DOL–CAR–GAP* model as an example, for each external test asset $r_{k,t}$, where $k = 1, \ldots, K$, we compare the marginal likelihoods of two competing regression specifications. The unrestricted model, denoted $\mathcal{M}_1$, allows for a non-zero pricing error (alpha) and is given by

$$\mathcal{M}_1: \quad r_{k,t} = \alpha_k + \beta_{k,\text{DOL}}\,\text{DOL}_t + \beta_{k,\text{CAR}}\,\text{CAR}_t + \beta_{k,\text{GAP}}\,\text{GAP}_t + \varepsilon_{k,t}, \tag{17}$$

while the competing restricted model, denoted $\mathcal{M}_0$, imposes the asset pricing restriction $\alpha_k = 0$:

$$\mathcal{M}_0: \quad r_{k,t} = \beta_{k,\text{DOL}}\,\text{DOL}_t + \beta_{k,\text{CAR}}\,\text{CAR}_t + \beta_{k,\text{GAP}}\,\text{GAP}_t + \varepsilon_{k,t}. \tag{18}$$

For each test asset we compute the Bayes factor BF_{01} as the ratio of the marginal likelihood of the restricted model to that of the unrestricted model, following the definition in Section 3.4. Evidence in favor of $\mathcal{M}_0$ (i.e., a higher marginal likelihood for the restricted model) implies that the *DOL–CAR–GAP* model successfully spans the test asset $r_{k,t}$.

Given the large cross-section of $K = 34$ currency portfolios, we summarize the results by reporting, for each specification, the percentage of test assets for which spanning is supported under both error distributions. Table 7 reports these spanning percentages for the full sample of monthly observations from January 1991 to March 2024, using the same in-sample training window as in the model scan.

Under the Gaussian specification, the *DOL–CAR–GAP* benchmark spans about 85% of the test assets, and this fraction rises to roughly 94% under the Student-t specification. The alternative specification *DDOL–CAR–GAP* delivers very similar spanning rates: around 76% under Gaussian errors and 94% under Student-t errors, reinforcing the interpretation of a market–carry–output gap structure. The remaining specifications in the table (*DOL–CAR*, *DB–CAR*, *DOL–CAR–MOM*, *DOL–MOM*, *DOL–VAL*, *DOL–IMB*, and *DOL–CAR–IMB*) serve as competing models. None of these alternative models dominate the *DOL–CAR–GAP* and *DDOL–CAR–GAP* structures uniformly across both Gaussian and Student-t error specifications. Overall, the evidence points to a robust role for a market proxy, carry, and the output gap in jointly pricing the cross-section of currency portfolios.

4.1.2 Frequentist Pricing Test

To evaluate the asset pricing performance from a frequentist perspective, we employ standard time-series regressions estimated by ordinary least squares (OLS). In contrast to the Bayesian framework in Section 4.1.1, which analyzes the models under both Gaussian and Student–t specifications, this benchmark analysis assumes that the regression errors follow a Gaussian distribution. Moreover, because the frequentist analysis does not rely on a separate training sample for prior construction, it lends itself naturally to direct subsample estimation. We therefore conduct a series of subsample analyses across alternative observation horizons to assess the temporal stability of pricing performance and to examine whether the leading specification identified in the full sample remains robust across subsamples.

Taking the *DOL–CAR–GAP* model as an illustrative example, for each test portfolio k, $k = 1, \ldots, K$, we estimate the following factor pricing regression:

$$r_{k,t} = \alpha_k + \beta_{k,\mathrm{DOL}}\,\mathrm{DOL}_t + \beta_{k,\mathrm{CAR}}\,\mathrm{CAR}_t + \beta_{k,\mathrm{GAP}}\,\mathrm{GAP}_t + u_{k,t}, \tag{19}$$

where $r_{k,t}$ denotes the excess return of portfolio k at time t, and the regression disturbance is assumed to be Gaussian, $u_{k,t} \sim \mathcal{N}(0, \sigma_k^2)$. The intercept α_k captures the pricing error, while the slope coefficients measure the portfolio's exposure to systematic risk factors.

To evaluate the overall economic magnitude of pricing errors and the explanatory power of the model across the cross section of $K = 34$ currency portfolios, we report the following cross-sectional average performance measures:

Average Absolute Alpha ($\overline{|\alpha|}$): The cross-sectional mean of the absolute pricing errors. This metric summarizes the average magnitude of unexplained returns across test assets; smaller values reflect better pricing performance.

Average Alpha Standard Error ($\overline{\mathrm{SE}(\alpha)}$): The cross-sectional mean of the standard errors of the estimated pricing errors. This measure captures the typical statistical precision with which alphas are estimated; smaller values indicate more precise inference.

Relative Absolute Alpha ($\overline{|\alpha|}/\overline{|\bar{r}|}$): The ratio of the average absolute pricing error to the average absolute mean return across test assets. This adjusts the magnitude of mispricing relative to the economic scale of typical returns.

Relative Squared Alpha ($\overline{\alpha^2}/\overline{\bar{r}^2}$): The ratio of the average squared pricing error to the average squared mean return across assets. By penalizing large pricing errors quadratically, this metric places greater weight on severe mispricing.

RMSE of Pricing Errors ($\overline{RMSE_\alpha}$): A root-mean-square measure of pricing errors constructed as

$$\overline{RMSE_\alpha} = \sqrt{\overline{\alpha^2} + \overline{\mathrm{SE}(\alpha)^2}}.$$

This statistic aggregates both the cross-sectional magnitude of pricing errors and their estimation uncertainty, and summarizes the overall scale of pricing deviations implied by the model.

Adjusted R^2 ($\overline{R^2_{\text{adj}}}$): The cross-sectional average of the adjusted R^2 from the time-series regressions, reported in percentage terms. A higher value indicates stronger time-series explanatory power of the factor model.

GRS Test p-Value (GRS$_p$): The p-value of the GRS test for the joint null hypothesis that all pricing errors are zero, $H_0 : \alpha_1 = \cdots = \alpha_K = 0$. A small p-value indicates rejection of the model's ability to jointly price all test assets.

Table 8 reports pricing performance across all specifications using the full sample of $K = 34$ test assets over the period January 1991 to March 2024 at the monthly frequency. Comparing with other factor models, the *DOL–CAR–GAP* model delivers the strongest overall pricing performance across all measures related to pricing errors. It achieves the smallest average absolute alpha ($\overline{|\alpha|}$) and the lowest pricing-error RMSE ($\overline{RMSE_\alpha}$), indicating that both the economic magnitude of mispricing and its estimation uncertainty are minimized under this specification. Although ranked second smallest, its average alpha standard error ($\overline{\text{SE}(\alpha)}$) of 0.0588 is very close to the lowest value of 0.0581. The relative absolute and relative squared alpha measures ($\overline{|\alpha|}/\overline{|\bar{r}|}$ and $\overline{\alpha^2}/\overline{\bar{r}^2}$) are also the lowest among all competing models, confirming that mispricing is small not only in absolute terms but also relative to the scale and variability of currency excess returns. In addition, the *DOL–CAR–GAP* model attains the highest average adjusted R^2 values of around 70.25%, reflecting strong time-series goodness-of-fit across the 34 test portfolios.

From a joint pricing perspective, the GRS test rejects the null hypothesis of zero pricing errors at the 5% significance level for all competing specifications with p-values below 0.05. The only exceptions are the *DOL–CAR–GAP* model, its economically motivated analogue *DDOL–CAR–GAP*, and the *DB–CAR* model, for which the GRS p-values exceed 0.05. While the *DB–CAR* model is still rejected at the 10% significance level, the *DOL–CAR–GAP* model achieves the highest GRS p-value of 0.2932. These results indicate that joint pricing is statistically rejected for most alternative models, whereas the dollar–carry–output gap structure remains the only specification that cannot be rejected at conventional significance levels within the frequentist framework.

Table 9 further reports the corresponding frequentist pricing results across three subsamples to assess the temporal stability of these conclusions. Across all three subsamples, the *DOL–CAR–GAP* model continues to exhibit consistently strong performance relative to competing specifications. In the front and middle subsamples reported in Panels A and B, it delivers the smallest or near-smallest average absolute alpha ($\overline{|\alpha|}$) and pricing-error RMSE ($\overline{RMSE_\alpha}$) across all candidate models, together with the highest adjusted R^2 values, exceeding 69% and 75%, respectively. In the end subsample reported in Panel C, although overall pricing errors decline for most specifications, the *DOL–CAR–GAP* and *DOL–CAR–IMB* models remain the dominant specifications across all performance measures, with similar magnitudes. The GRS test also reveals a clear temporal pattern: most models are strongly rejected in the earlier subsamples, whereas none of the specifications are rejected in the most recent subsample. Importantly, across all subperiods, the *DOL–CAR–GAP* model attains the highest or near-highest GRS p-values, remaining consistently farther from rejection than alternative models.

Overall, the subsample evidence confirms that the superior pricing performance of the dollar–carry–output gap structure is not driven by any particular period but remains economically and statistically robust across distinct market regimes.

4.2 Investment Performance

4.2.1 Minimum-Variance Frontiers

While the pricing tests focus on the ability of factor models to explain the cross section of currency returns, investors ultimately care about the risk–return trade-offs that these factors deliver in portfolio space. To evaluate the economic value of each model from a portfolio allocation perspective, we therefore complement the pricing results with a comparison of their implied minimum-variance frontiers (MVFs). The MVF characterizes the full set of efficient portfolios attainable from a given factor set and provides a direct assessment of the maximum investment opportunities available to minimum-variance investors. By examining how the frontiers differ across competing specifications, we can assess not only whether a model prices well statistically, but also whether it offers superior diversification benefits and attainable Sharpe ratios in economic terms.

Figure 4 reports the Markowitz minimum-variance frontiers implied by the Bayesian best model and a set of competing currency factor specifications, where *DOL* and *DDOL* serve as alternative proxies for the average market factor. The frontiers are constructed from the sample moments (sample means and covariance matrices) of the monthly returns on the candidate currency factors over the period from January 1991 to March 2024. For each model, we compute the set of minimum-variance portfolios that can be formed from its corresponding factor set. In the figure, the dot on each curve marks the global minimum–variance portfolio, the solid segment denotes the efficient portion of the frontier (portfolios with higher expected returns than the minimum–variance portfolio), and the dashed segment denotes the inefficient portion.

The figure reveals economically meaningful differences in portfolio investment opportunities across models. The Bayesian best specification, the *DOL–CAR–GAP* three-factor model, consistently traces out the upper envelope of the frontiers and spans the widest range of attainable risk–return combinations, extending toward both lower-risk and higher-return regions relative to the competing specifications. The alternative global factor specification *DDOL–CAR–GAP* delivers a frontier that lies close to that of *DOL–CAR–GAP* over the central range of volatilities, indicating that both *DOL* and *DDOL* are viable proxies for the global average factor. However, the *DDOL–CAR–GAP* frontier is somewhat more compressed, covering a narrower range of risk–return trade-offs than its *DOL*-based counterpart.

Models such as *DOL–CAR*, *DOL–CAR–MOM*, and *DOL–CAR–IMB* generate frontiers that remain relatively close to the benchmark but are uniformly dominated by *DOL–CAR–GAP* and *DDOL–CAR–GAP*. By contrast, simpler two-factor specifications like *DOL–MOM*, *DOL–VAL* and *DOL–IMB* trace out noticeably lower frontiers, indicating that they fail to fully span the cross-section of currency portfolio returns. The *DB–CAR* specification provides an interesting contrast: its efficient frontier becomes relevant only at relatively high volatility levels, emerging to the right of the other curves. Although it attains comparatively high expected returns once it becomes feasible, these outcomes are achieved only at substantially elevated risk levels.

Overall, the minimum-variance comparison confirms that the *DOL–CAR–GAP* model and its *DDOL–CAR–GAP* analogue not only price currency portfolios well, but also offer the most attractive risk–return opportunities for a minimum-variance investor over the sample period.

4.2.2 Out-of-Sample Sharpe Ratios

To evaluate the real-time investment performance of alternative factor models, we conduct an out-of-sample (OOS) portfolio exercise in which trading strategies are updated recursively at each point in time, following the real-time portfolio evaluation framework of Chib et al. (2024). At the end of every month, we re-estimate the optimal tangency portfolio using only the information available up to that date, exactly as a real investor would. These newly updated portfolio weights are then implemented in the following month to form an ex ante trading strategy. The resulting realized excess return is recorded, after which the information set is expanded to include this new observation and the strategy is updated again. This recursive procedure generates a sequence of realized portfolio returns based entirely on real-time information, rather than ex post full-sample estimates.

The economic performance of each model is evaluated using the annualized Sharpe ratio computed from these realized OOS returns. Because portfolio weights are continuously re-optimized as new data arrive, this metric captures not only each model's average return and risk, but also its ability to adapt dynamically to evolving market conditions. In this sense, the annualized OOS Sharpe ratio provides a strictly forward-looking and practically relevant measure of investment performance, directly reflecting the gains and risks that an investor would have experienced when implementing the trading strategy in real time.

Table 10 reports the one-year out-of-sample annualized Sharpe ratios for the Bayesian best model and several competing currency factor models under both Gaussian and Student-t return specifications. The in-sample and training samples used in these evaluations are aligned with the setup described in Section 3.2. Under the Gaussian specification, the *DOL–CAR–GAP* model remains among the top-performing specifications, delivering an OOS Sharpe ratio of 2.4073. Consistent with this finding, under the Student-t assumption, the Bayesian best *DOL–CAR–GAP* three-factor model delivers the strongest out-of-sample performance among all competing models, achieving the highest annualized Sharpe ratio of **2.9950**. This result confirms that the model selected by the Bayesian framework in-sample also provides the most robust real-time investment performance out of sample. Other leading specifications, such as *DDOL–CAR–GAP*, *DOL–CAR*, and *DOL–CAR–IMB*, also exhibit strong but slightly weaker performance, with Student-t OOS Sharpe ratios of 2.7401, 2.6792, and 2.6973, respectively.

Overall, the out-of-sample evidence strongly reinforces the in-sample model comparison results. The Bayesian best *DOL–CAR–GAP* model consistently dominates alternative well-documented currency factor models in terms of real-time risk-adjusted performance. Its superior and stable OOS Sharpe ratios demonstrate that this factor combination not only fits the data well in sample, but also delivers economically meaningful and robust portfolio gains in realistic trading environments.

Beyond model ranking, the OOS portfolio results also provide complementary economic evidence in favor of heavy-tailed return dynamics. While the *DOL–CAR–GAP* specification remains the best-performing model under both Gaussian and Student-t assumptions, its dominance is most pronounced under the Student-t framework, where it achieves the highest OOS Sharpe ratio. This pattern indicates that explicitly accounting for fat-tailed risks is not only statistically relevant, but also economically beneficial from an investor's perspective. From a portfolio allocation standpoint, allowing for heavy tails directly affects the assessment of downside risk and extreme return realizations, which in turn shapes optimal tangency portfolio weights. The superior OOS performance under the Student-t specification suggests that investors who properly account for tail risk are better compensated in real time, particularly during periods of elevated market stress and sudden shifts in global risk sentiment. In contrast, Gaussian-based portfolios tend to understate tail risk and therefore deliver inferior risk-adjusted performance when exposed to extreme currency movements. Taken together with the earlier marginal likelihood evidence, these OOS investment results reinforce the conclusion that heavy-tailed distributions provide a more realistic and economically meaningful description of currency factor returns and that the dominance of the *DOL–CAR–GAP* model is, if anything, understated under the Gaussian benchmark.

Summary of the Economic Performance of *DOL–CAR–GAP*

In summary, our economic performance evaluation shows that the *DOL–CAR–GAP* model turns its statistical strength into solid economic performance. On the pricing side, the Bayesian spanning tests show that the dollar–carry–output gap structure spans most external currency portfolios. The frequentist GRS test leads to the same conclusion for both *DOL–CAR–GAP* and *DDOL–CAR–GAP*, as neither model is rejected. When we examine pricing errors and goodness of fit, the frequentist regressions clearly identify *DOL–CAR–GAP* as the best-performing model specification.

On the investment side, the *DOL–CAR–GAP* model delivers the most attractive risk–return trade-offs in the minimum-variance frontier. Under both Gaussian and Student-t assumptions, it achieves the highest out-of-sample annualized Sharpe ratios, with much stronger performance under the Student-t specification. This pattern highlights the economic value of modeling heavy-tailed return dynamics and shows that the Student-t specification is preferred to the Gaussian benchmark.

5 Conclusion

This paper views the selection of currency asset-pricing models as a *survival-of-the-fittest* exercise. Building on Chib and Zeng (2020) and Chib et al. (2020), we employ a Bayesian model scan to evaluate combinations of all economically motivated currency risk factors, allowing the data to identify the strongest performers rather than testing a single prespecified model. This approach establishes a data-driven criterion based on posterior model probabilities and marginal likelihoods, jointly analyzing factor dynamics and no-arbitrage conditions to determine the optimal stochastic discount factor without requiring exogenously given test assets.

Our analysis identifies the parsimonious three-factor model *DOL–CAR–GAP* as the dominant pricing kernel. This specification attains the highest marginal likelihoods and concentrates the posterior mass under both Gaussian and heavy-tailed Student-t assumptions. Importantly, the model evidence favors the Student-t specification, underscoring the relevance of heavy-tailed return dynamics in currency markets. These results are robust to alternative degrees of freedom and to replacing *DOL* with *DDOL* as the market factor proxy. Posterior analysis further indicates that the three factors command positive risk premia and effectively span the remaining factors (i.e., *MOM*, *DB*, *VAL*, *FXC*, and *IMB*). The out-of-sample predictive performance corroborates these in-sample findings.

These statistical results translate into economically meaningful gains. We evaluate the *DOL–CAR–GAP* model from both pricing and investment perspectives. Using an exogenously given set of currency portfolios as test assets, we assess pricing performance through Bayesian model comparison and standard frequentist regressions. While Bayesian analysis and GRS tests confirm the spanning ability of both *DOL–CAR–GAP* and its analogue *DDOL–CAR–GAP*, traditional frequentist metrics also clearly favor *DOL–CAR–GAP*, reflecting lower pricing errors and superior goodness of fit. From an investment perspective, the minimum–variance frontier supports the efficiency of the model. Moreover, real-time out-of-sample investment exercises show that *DOL–CAR–GAP* delivers the highest annualized Sharpe ratios, particularly under the Student-t specification, highlighting the economic importance of modeling tail risk in currency returns.

Methodologically, our findings validate the Bayesian model scan as a scalable and effective approach to model selection in the presence of a large set of candidate factors. By allowing for flexible distributional assumptions and avoiding reliance on prespecified test assets, the approach identifies the *survival-of-the-fittest* specification in a disciplined and data-driven manner. While our analysis focuses on currency markets, the methodology provides a versatile blueprint that can be applied to other asset classes, including equities, commodities, and cryptocurrencies.

Ultimately, both the empirical evidence and the economic performance evaluation point to a unified and economically interpretable benchmark. Despite the proliferation of proposed currency factors, we show that risk compensation is best captured by a coherent structure centered on the dollar (*DOL*), the carry trade (*CAR*), and the output gap (*GAP*) which reflects the business cycle. The *DOL–CAR–GAP* model therefore provides a natural reference for currency risk compensation and a robust foundation for future research in international asset pricing.

Tables and Figures

Table 1 Definitions and Construction of Candidate Currency Factors

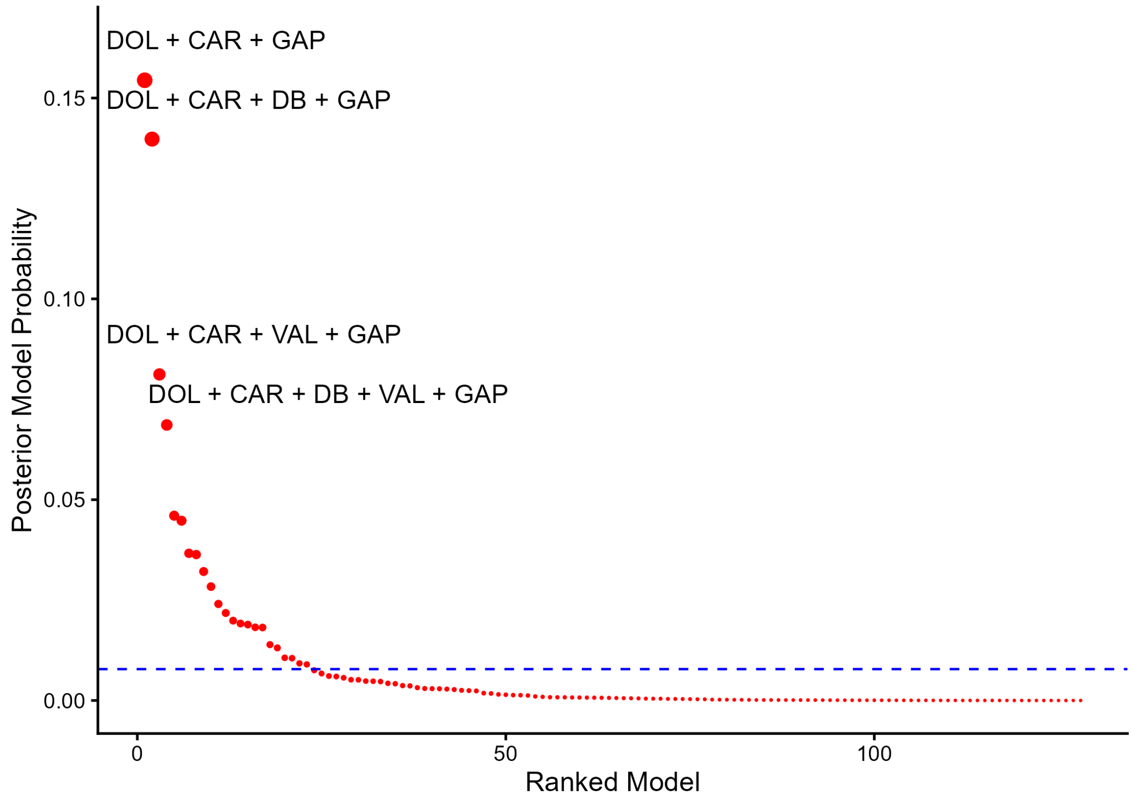
Factor	Description	Construction	Reference
DOL	Dollar	Long all foreign currencies and short the US dollar.	Lustig et al. (2011)
DDOL	Dollar Carry	Long (Short) DOL when the median forward discount is positive (negative).	Lustig et al. (2014)
CAR	Carry	Long high-interest-rate currencies and short low-interest-rate currencies.	Lustig et al. (2011)
MOM	Momentum	Long high-past-return currencies and short low-past-return currencies.	Menkhoff et al. (2012b)
DB	Dollar Beta	Long (Short) currencies with high DOL loadings and short (long) currencies with low DOL loadings when the median forward discount is positive (negative).	Verdelhan (2018)
VAL	Value	Long under-valued currencies and short over-valued currencies.	Menkhoff et al. (2017)
FXC	FX Correlation	Long currencies with high loadings to FX correlation dispersion and short currencies with low loadings to FX correlation dispersion.	Mueller et al. (2017)
GAP	Output Gap	Long high-output gap currencies and short low-output gap currencies.	Colacito et al. (2020)
IMB	Global Imbalance	Long risky currencies and short safe currencies.	Corte et al. (2016)

Note: This table summarizes the definitions and construction methodologies for the nine currency factors included in our benchmark model scan. For each factor, we report the standard abbreviation, a descriptive label, the specific portfolio construction strategy, and the literature reference.

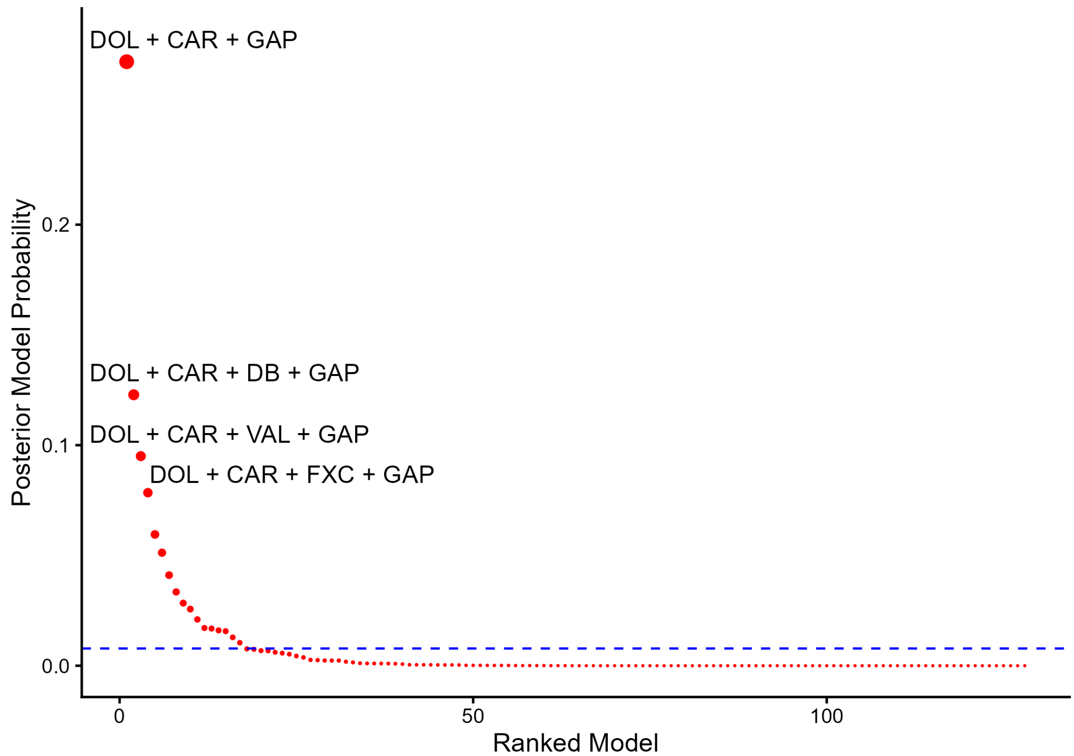
Table 2 Descriptive Statistics of Candidate Currency Factors

Factors	Mean	Std Dev.	Skewness	Kurtosis	Max	Min
DOL	0.326	24.110	-0.491	4.349	65.217	-100.217
DDOL	2.999	23.924	-0.231	4.480	100.217	-93.653
CAR	4.657	25.540	-1.049	4.999	55.106	-100.696
MOM	1.860	45.055	-0.386	7.478	198.633	-276.491
DB	4.507	34.883	-0.228	3.341	102.880	-110.807
VAL	1.428	23.528	0.267	3.916	92.715	-84.065
FXC	0.216	23.779	0.041	4.064	82.390	-78.940
GAP	4.140	24.882	-0.130	6.711	102.670	-124.733
IMB	1.951	22.927	-0.713	5.955	77.797	-112.594

Note: This table reports the annualized mean (**Mean**), standard deviation (**Std Dev.**), skewness (**Skewness**), kurtosis (**Kurtosis**), maximum (**Max**) and minimum (**Min**) values for each currency factor. Numbers are in percentage. The sample spans from January 1991 to March 2024.



(a) Gaussian Specification



(b) Student-t Specification

Figure 1 Posterior Model Probabilities for Ranked Models under Gaussian and Student-t Specifications with *DOL* Restriction

Table 3 Marginal Likelihoods of Selected Currency Factor Models

Model	Logmarg	
	Gaussian	**Student-t**
DOL CAR GAP (Bayesian best model)	**6408.038**	**6560.042**
DOL CAR	6406.468	6556.236
DB CAR	6407.150	6556.989
DOL CAR MOM	6404.426	6554.439
DOL MOM	6399.737	6537.475
DOL VAL	6399.942	6537.380
DOL IMB	6401.793	6541.798
DOL CAR IMB	6404.641	6554.412

Note: This table reports the log–marginal likelihoods for the Bayesian best model and a set of representative competing currency factor models from the literature under both Gaussian and Student–t specifications. The sample spans January 1991 to April 2023. The first 15% of the data (January 1991– October 1995) is used as the training sample, while the remaining 329 monthly observations (November 1995–April 2023) constitute the estimation sample.

Table 4 Marginal Likelihood of the *DOL–CAR–GAP* Model Across Degrees of Freedom

ν	Logmarg
2.1	6445.009
3	6544.252
3.5	6554.089
4	6558.470
4.5	**6560.099**
5	**6560.042**
5.5	6558.973
6	6557.214
6.5	6555.059
7	6552.603
8	6547.216
9	6541.545
10	6535.898
20	6490.218
30	6462.160
40	6443.751
50	6430.795
60	6421.196
∞	6408.038

Note: This table reports the log marginal likelihoods of the *DOL–CAR–GAP* three-factor model under Student-t return distributions with different degrees of freedom ν. The case $\nu = \infty$ corresponds to the Gaussian specification. The highest log marginal likelihoods occur around $\nu = 4.5$ and $\nu = 5$ (in bold), supporting the baseline choice of $\nu = 5$ while indicating that model fit is robust over a broad range of tail thickness.

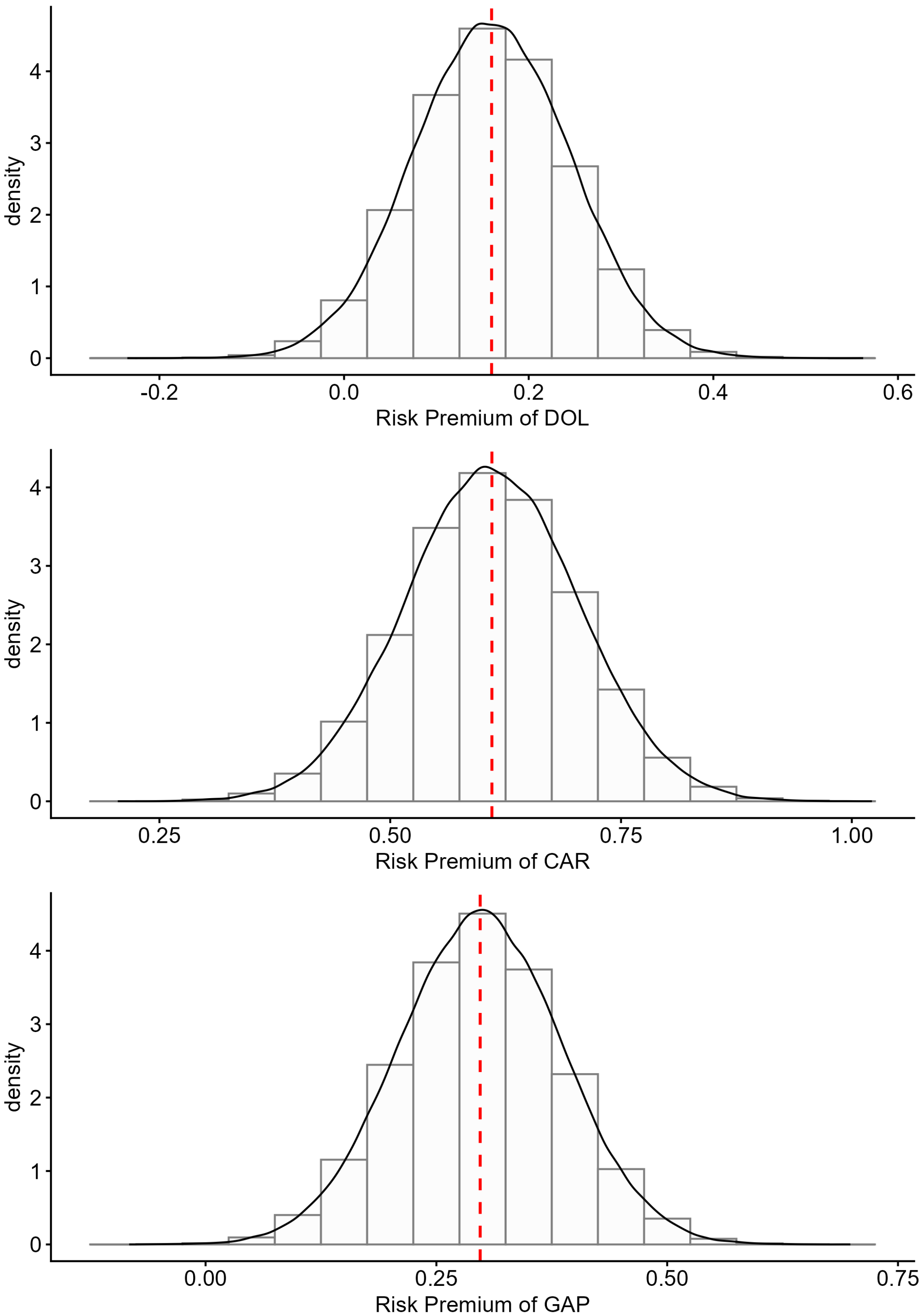


Figure 2 Posterior Risk Premia Distributions for *DOL*, *CAR*, and *GAP*
Note: This figure displays the posterior distributions of the risk premia for the three factors *DOL*, *CAR* and *GAP* in the best model under the Student-t specification with $\nu = 5$. The dashed red lines indicate posterior means.

Table 5 Bayesian Spanning Test Results of Losing Factors by the *DOL–CAR–GAP* Three-Factor Model

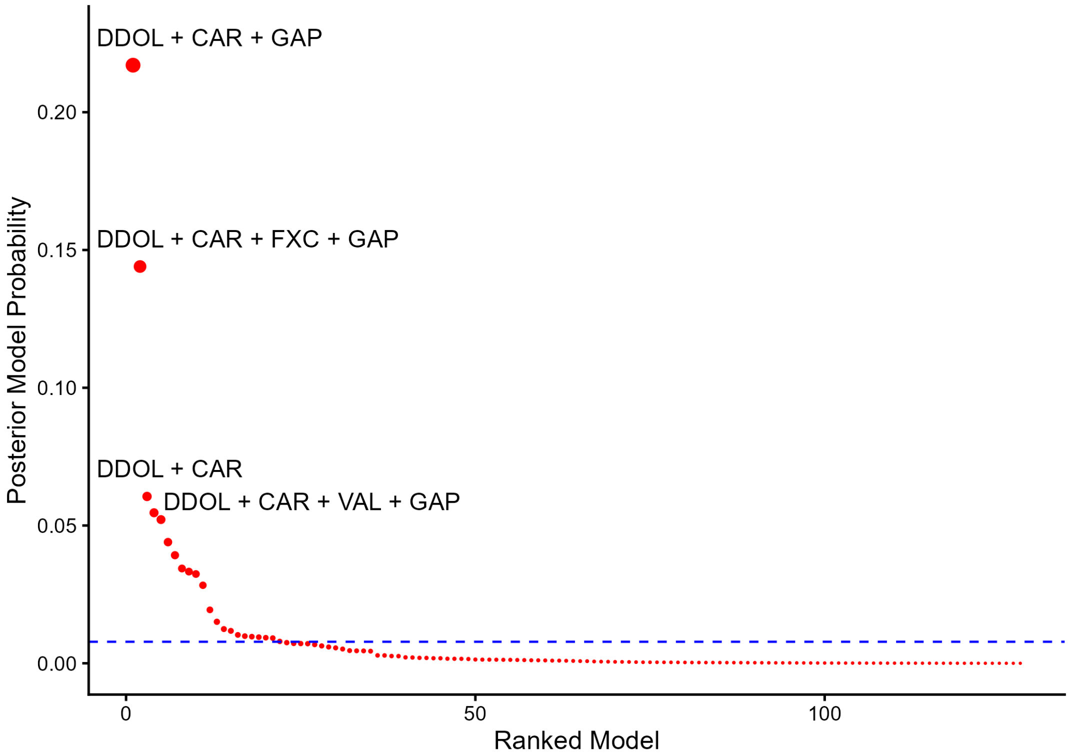
Losing Factor	Logmarg ($\mathcal{M}_0$)	Logmarg ($\mathcal{M}_1$)	$\ln(\mathrm{BF}_{01})$	BF_{01}
	(Restricted)	(Unrestricted)		
MOM	626.6843	624.7399	1.9444	6.9895
DB	716.2017	715.6674	0.5343	1.7062
VAL	849.9426	849.2896	0.6530	1.9214
FXC	845.1698	844.8657	0.3041	1.3555
IMB	901.9140	901.0756	0.8383	2.3125

Note: This table reports log marginal likelihoods for two competing regression specifications for each losing factor, where the factor's excess return is projected on the *DOL–CAR–GAP* three-factor model. The restricted model $\mathcal{M}_0$ imposes zero intercept (exact spanning), while the unrestricted model $\mathcal{M}_1$ allows for a nonzero intercept. BF_{01} denotes the Bayes factor in favor of the restricted model. Values of $\mathrm{BF}_{01} > 1$ (equivalently, $\ln(\mathrm{BF}_{01}) > 0$) indicate that the data favor the hypothesis that the losing factor is fully spanned by the *DOL–CAR–GAP* pricing kernel.

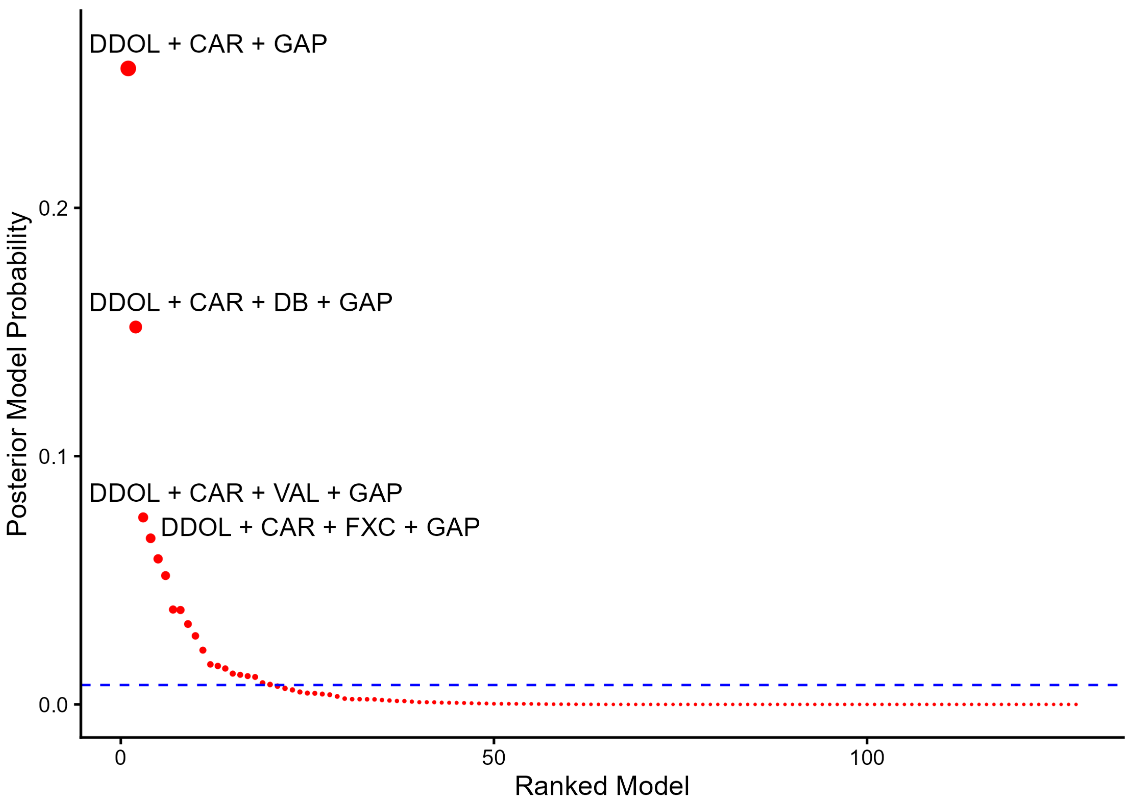
Table 6 Out-of-sample Predictive Likelihoods of Selected Currency Factor Models

Model	**Log Predicted Likelihood**	
	Gaussian	**Student-t**
DOL CAR GAP (Bayesian best model)	**255.468**	**259.474**
DOL CAR	255.170	258.651
DB CAR	254.807	258.534
DOL CAR MOM	255.134	258.449
DOL MOM	254.163	256.068
DOL VAL	254.090	256.202
DOL IMB	254.530	256.927
DOL CAR IMB	255.160	258.659

Note: This table reports out-of-sample (OOS) log predictive likelihoods for the Bayesian best model and a set of representative competing currency factor models from the literature under both Gaussian and Student-*t* specifications. The out-of-sample uses data from the last 12 months in the full sample: April 2023 to March 2024.



(a) Gaussian Specification



(b) Student-t Specification

Figure 3 Posterior Model Probabilities for Ranked Models under Gaussian and Student-t Specifications with *DDOL* Restriction

Table 7 Percentage of Test Assets Spanned by Selected Currency Factor Models: Full Sample

Model	**Percentage Spanned** $(\mathbf{Pct}_{\mathcal{M}_0 > \mathcal{M}_1})$	
	Gaussian	**Student-t**
DOL CAR GAP	**85.29**	**94.12**
DDOL CAR GAP	**76.47**	**94.12**
DOL CAR	76.47	88.24
DB CAR	73.53	97.06
DOL CAR MOM	70.59	79.41
DOL MOM	73.53	67.65
DOL VAL	73.53	67.65
DOL IMB	64.71	67.65
DOL CAR IMB	76.47	88.24

Note: This table reports, for each currency factor model, the percentage of test assets that are spanned, defined as the fraction of test assets for which the intercept-free model $\mathcal{M}_0$ is favored over the model with an intercept $\mathcal{M}_1$ in the Bayesian model comparison. The full sample covers monthly observations from January 1991 to March 2024, and the training sample coincides with the in-sample training window used in the model scan. We report these percentages separately under Gaussian and Student-t specifications for the error terms.

Table 8 Frequentist Pricing Test Performance of Selected Currency Factor Models: Full Sample

Model	$\overline{\lvert\alpha\rvert}$	$\overline{\text{SE}(\alpha)}$	$\overline{\lvert\alpha\rvert}/\overline{\lvert\bar{r}\rvert}$	$\overline{\alpha^2}/\overline{\bar{r}^2}$	$\overline{RMSE_\alpha}$	$\overline{R^2_{\text{adj}}}$	GRS_p
DOL CAR GAP	**0.0623**	**0.0588**	**0.5969**	**0.3672**	**0.1092**	**70.2498**	**0.2932**
DDOL CAR GAP	0.1223	0.0970	1.1720	0.9180	0.1706	28.3136	**0.2884**
DOL CAR	0.0803	0.0599	0.7697	0.6375	0.1330	68.8839	0.0472
DB CAR	0.1324	0.1002	1.2684	1.1273	0.1844	22.1439	0.0611
DOL CAR MOM	0.0777	0.0592	0.7442	0.5967	0.1295	69.4627	0.0187
DOL MOM	0.0959	0.0609	0.9190	0.8576	0.1494	67.6072	0.0007
DOL VAL	0.1044	0.0605	1.0002	1.0010	0.1590	68.0725	0.0009
DOL IMB	0.1014	0.0594	0.9711	0.9080	0.1526	68.7123	0.0045
DOL CAR IMB	0.0807	**0.0581**	0.7733	0.6431	0.1328	70.2184	0.0485

Note: This table reports frequentist pricing test results for selected currency factor models using the full sample of $K = 34$ test assets using monthly observations from January 1991 to March 2024. $\overline{\lvert\alpha\rvert}$, $\overline{\text{SE}(\alpha)}$, $\overline{RMSE_\alpha}$, and $\overline{R^2_{\text{adj}}}$ are reported in percentage terms. $\overline{\lvert\alpha\rvert}/\overline{\lvert\bar{r}\rvert}$ and $\overline{\alpha^2}/\overline{\bar{r}^2}$ are unit-free ratios. GRS_p denotes the p-value of the GRS test for the joint null hypothesis that all pricing errors are zero.

Table 9 Frequentist Pricing Test Performance of Selected Currency Factor Models: Subsamples

Model	$\overline{\lvert\alpha\rvert}$	$\overline{\text{SE}(\alpha)}$	$\overline{\lvert\alpha\rvert}/\overline{\lvert\bar{r}\rvert}$	$\overline{\alpha^2}/\overline{\bar{r}^2}$	$\overline{RMSE_\alpha}$	$\overline{R^2_{\text{adj}}}$	GRS$_p$
Panel A: Front Subsample (January 1991 – December 2005)							
DOL CAR GAP	**0.1063**	**0.0900**	**0.5281**	**0.2840**	**0.1753**	**69.2043**	**0.2008**
DDOL CAR GAP	0.1539	0.1443	0.7645	0.4691	0.2387	29.3204	**0.2480**
DOL CAR	0.1612	0.0909	0.8005	0.7384	0.2548	66.3361	0.0114
DB CAR	0.1732	0.1472	0.8600	0.5667	0.2549	22.1570	0.0288
DOL CAR MOM	0.1451	0.0902	0.7206	0.6054	0.2341	67.2089	0.0103
DOL MOM	0.1623	0.0934	0.8060	0.7083	0.2509	64.2448	0.0017
DOL VAL	0.1896	0.0920	0.9418	1.0506	0.2969	65.4001	0.0010
DOL IMB	0.1806	0.0887	0.8972	0.9482	0.2831	66.6203	0.0033
DOL CAR IMB	0.1588	0.0867	0.7889	0.7139	0.2496	68.7455	0.0125
Panel B: Middle Subsample (February 2000 – January 2015)							
DOL CAR GAP	**0.0832**	**0.0811**	**0.4459**	**0.2491**	**0.1431**	**75.2590**	**0.2572**
DDOL CAR GAP	0.1027	0.1330	0.5502	0.3201	0.1853	43.5486	**0.2375**
DOL CAR	0.0940	0.0828	0.5034	0.2927	0.1510	74.3326	0.1294
DB CAR	0.1102	0.1386	0.5903	0.3788	0.1965	36.9657	0.1460
DOL CAR MOM	0.0923	0.0818	0.4947	0.2822	0.1487	74.8453	0.1162
DOL MOM	0.0983	0.0834	0.5269	0.3262	0.1559	73.5724	0.0363
DOL VAL	0.1029	0.0827	0.5513	0.3547	0.1602	73.7569	0.0399
DOL IMB	0.1024	0.0826	0.5488	0.3397	0.1581	74.0765	0.0601
DOL CAR IMB	0.0945	0.0812	0.5062	0.2943	0.1507	75.0925	0.1363
Panel C: End Subsample (April 2009 – March 2024)							
DOL CAR GAP	**0.0502**	**0.0755**	**0.6391**	**0.3651**	**0.1052**	**75.3115**	**0.9009**
DDOL CAR GAP	0.1483	0.1338	1.8885	2.7231	0.2163	33.7035	**0.8101**
DOL CAR	0.0499	0.0775	0.6355	0.3615	0.1062	74.6241	0.8951
DB CAR	0.1712	0.1377	2.1805	3.5509	0.2367	29.6608	0.8489
DOL CAR MOM	0.0501	0.0769	0.6382	0.3835	0.1068	74.9934	0.8802
DOL MOM	0.0772	0.0792	0.9826	1.0002	0.1339	72.3959	0.4917
DOL VAL	0.0747	0.0771	0.9510	0.8987	0.1290	73.6300	0.4870
DOL IMB	0.0772	0.0778	0.9828	0.9227	0.1302	73.4334	0.5788
DOL CAR IMB	0.0492	0.0753	0.6262	0.3498	0.1041	75.9787	0.9162

Note: This table reports frequentist pricing test results for selected currency factor models across three subsamples. Panels A, B, and C correspond to the front, middle, and end subsamples, respectively. $\overline{\lvert\alpha\rvert}$, $\overline{\text{SE}(\alpha)}$, $\overline{RMSE_\alpha}$, and $\overline{R^2_{\text{adj}}}$ are reported in percentage terms. $\overline{\lvert\alpha\rvert}/\overline{\lvert\bar{r}\rvert}$ and $\overline{\alpha^2}/\overline{\bar{r}^2}$ are unit-free ratios. GRS$_p$ denotes the p-value of the GRS test for the joint null hypothesis that all pricing errors are zero.

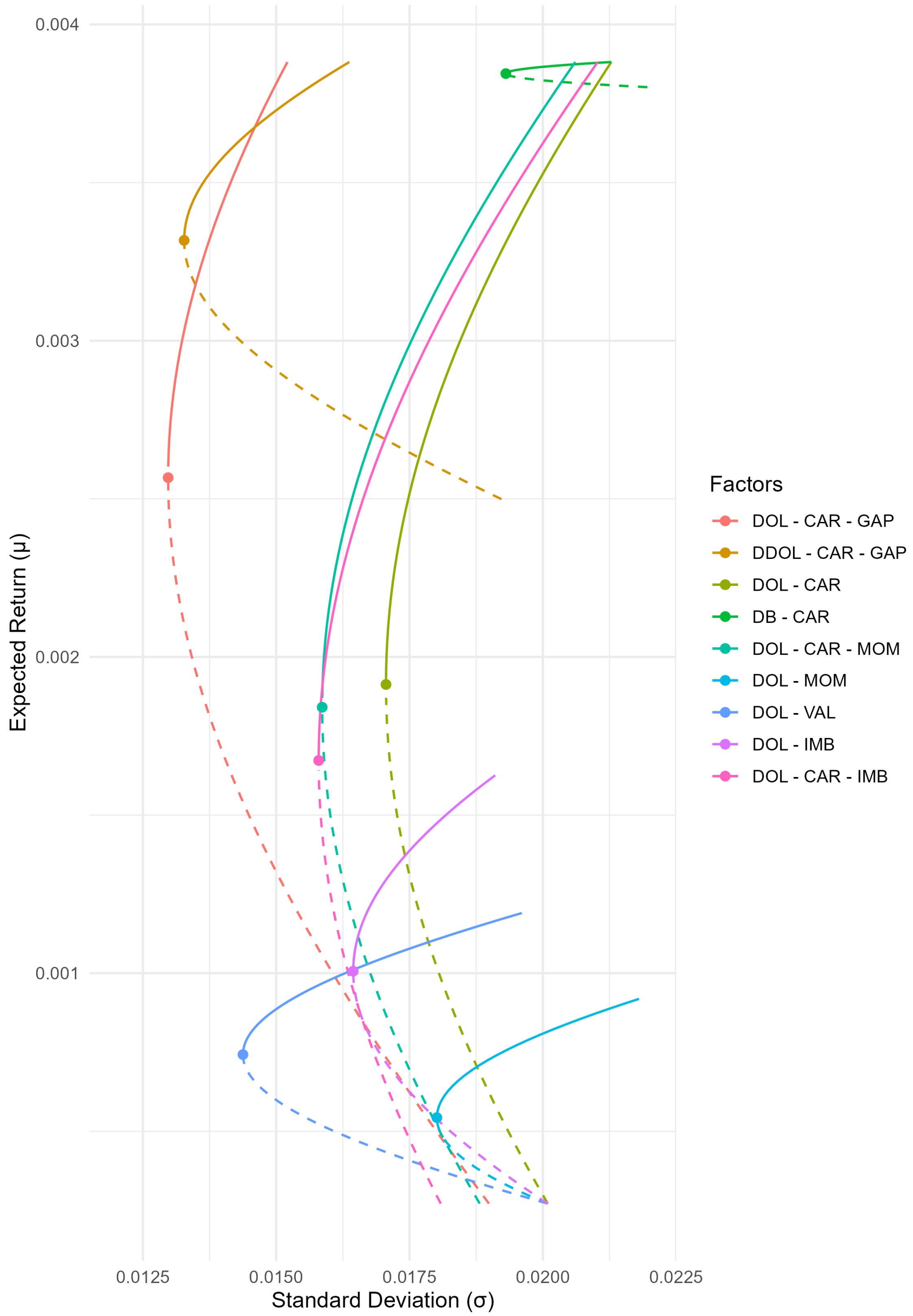


Figure 4 Minimum–Variance Frontiers

Note: This figure displays the Markowitz minimum-variance frontiers for portfolios formed from the factors in selected currency pricing models, using data from January 1991 to March 2024. For each model, the dot marks the global minimum–variance portfolio; solid lines show the efficient portion of the frontier (portfolios with higher expected returns than the global minimum–variance portfolio), while dashed lines show the inefficient portion.

Table 10 Out-of-Sample Annualized Sharpe Ratios for Selected Currency Factor Models

Model	Annualized SR_{OOS}	
	Gaussian	**Student-t**
DOL CAR GAP	2.4073	**2.9950**
DDOL CAR GAP	2.2792	2.7401
DOL CAR	2.2121	2.6792
DB CAR	1.6428	2.3151
DOL CAR MOM	2.1728	2.3229
DOL MOM	-0.1521	-0.5426
DOL VAL	-0.5876	-0.3208
DOL IMB	1.3137	1.1075
DOL CAR IMB	2.2263	2.6973

Note: This table reports the out-of-sample annualized Sharpe ratios based on recursively updated portfolio weights and one-month-ahead realized returns. The evaluation is conducted under both Gaussian and Student-t return specifications. The in-sample and training samples used for these evaluations are aligned with the setup described in Section 3.1. Sharpe ratios are annualized using monthly returns. Portfolio weights are determined based on recursively updated estimation.

References

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

Cespa, G., A. Gargano, S. J. Riddiough, and L. Sarno (2022). Foreign exchange volume. *The Review of Financial Studies 35*(5), 2386–2427.

Chernov, M., M. Dahlquist, and L. A. Lochstoer (2024). *An anatomy of currency strategies: The role of emerging markets*. National Bureau of Economic Research.

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the american statistical association 90*(432), 1313–1321.

Chib, S. and X. Zeng (2020). Which factors are risk factors in asset pricing? a model scan framework. *Journal of Business & Economic Statistics 38*(4), 771–783.

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

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

Colacito, R., S. J. Riddiough, and L. Sarno (2020). Business cycles and currency returns. *Journal of Financial Economics 137*(3), 659–678.

Corte, P. D., S. J. Riddiough, and L. Sarno (2016). Currency premia and global imbalances. *The Review of Financial Studies 29*(8), 2161–2193.

Corte, P. D., L. Sarno, M. Schmeling, and C. Wagner (2022). Exchange Rates and Sovereign Risk. *Management Science 68*, 5591–5617.

Dobrynskaya, V. (2014). Downside Market Risk of Carry Trades. *Review of Finance 18*(5), 1885–1913.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy 81*(3), 607–636.

Galsband, V. and T. Nitschka (2014). Currency Excess Returns and Global Downside Risk. *Journal of International Money and Finance 47*, 268–285.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica 57*(5), 1121–1152.

Lettau, M., M. Maggiori, and M. Weber (2014). Conditional Risk Premia in Currency Markets and Other Asset Classes. *Journal of Financial Economics 114*, 197–225.

Lustig, H., N. Roussanov, and A. Verdelhan (2011). Common Risk Factors in Currency Returns. *Review of Financial Studies 24*(11), 3731–3777.

Lustig, H., N. Roussanov, and A. Verdelhan (2014). Countercyclical Currency Risk Premia. *Jounal of Financial Economics 111*(3), 527–553.

Mancini, L., A. Ranaldo, and J. Wrampelmeyer (2013). Liquidity in the Foreign Exchange Market: Measurement, Commonality, and Risk Premiums. *Journal of Finance 68*(5), 1805–1841.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012a). Carry Trades and Global Foreign Exchange Volatility. *Journal of Finance 67*(2), 681–718.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012b). Currency Momentum Strategies. *Journal of Financial Economics 106*, 620–684.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2017). Currency Value. *Review of Financial Studies 30*(2), 416–441.

Mueller, P., A. Stathopoulos, and A. Vedolin (2017). International Correlation Risk. *Journal of Financial Economics 126*(2), 270–299.

Nucera, F., L. Sarno, and G. Zinna (2024). Currency risk premiums redux. *The Review of Financial Studies 37*(2), 356–408.

Qiao, Z., Y. Wang, and K. Lam (2022). New evidence on bayesian tests of global factor pricing models. *Journal of Empirical Finance 68*, 160–172.

Verdelhan, A. (2018). The share of systematic variation in bilateral exchange rates. *The Journal of Finance 73*(1), 375–418.