International Mispricing and Arbitrage Premia*

Mirela Sandulescu† Paul Schneider‡

November 2021

Abstract

We develop Residual MisPricing (RMP), an index capturing mispricing relative to a linear benchmark model for basic assets from the absence of arbitrage. We show that conditional linear asset pricing models perform well on average and during normal times, while they imply larger RMP during crisis periods. Return data for several economies reveal that RMP is countercyclical, related to financial uncertainty, and positively correlated with conditional international equity and currency risk premia. We find that RMP predicts future market dislocations, including covered interest rate parity deviations.

Keywords: stochastic discount factor, residual mispricing, financial uncertainty, exchange rates, economically constrained machine learning.

JEL Classification: G11, G12, G15

*We thank Rohit Allena, Andrada Bilan, Aref Bolandnazar, Oleg Bondarenko, Xavier Gabaix, Stefano Giglio, Andrey Malenko, Shane Miller, Paolo Pasquariello, Uday Rajan, Angelo Ranaldo, Shri Santosh, Christian Wagner, Mao Ye and seminar and conference participants at Michigan Ross, Goethe University, Yale-SoFiE Machine Learning Virtual Conference, AFFI Conference, FMA Annual Meeting and the 10th Workshop on Exchange Rates for helpful comments and discussions. All errors are our own.

†University of Michigan, Ross School of Business. Email: mirelas@umich.edu

‡USI Lugano and Swiss Finance Institute. Email: paul.schneider@usi.ch
1 Introduction

Linear asset pricing models are widely used in industry and academia for prediction, pricing, and benchmarking. While they can rule out statistical arbitrage in large cross sections, they imply the existence of arbitrage portfolios in smaller samples, at odds with basic financial theory. In this paper, we propose the smallest economically motivated modification to benchmark linear asset pricing models that is compatible with the no-arbitrage condition. We then define the difference between the linear model and its closest no-arbitrage counterpart as residual mispricing (RMP). In economic terms, it can be interpreted as the cheapest trading strategy insuring the linear asset pricing model against arbitrage. Importantly, data requirements of RMP are minimal, using only return data on basic assets such as equities and bonds. Empirically, we derive RMP for several developed markets along with their exchange rates. We find a highly significant relation to indicators of financial uncertainty, financial distress measures, conditional risk premia, as well as the cross section of currency returns.

We develop several methodological and empirical contributions to obtain our results. First, following Hansen and Jagannathan (1991, HJ) we are agnostic about the factor model driving the cross-section of international returns, and recognize that it must lie within the space of linear combinations of the returns available to investors. However, in its widely used linear form, the nonparametric HJ minimum-variance pricing framework does not rule out the existence of arbitrage portfolios. Its arbitrage-free version (the positive part of a linear pricing model) exhibits a type of nonlinearity that is difficult to estimate, in particular conditionally, and therefore rarely used in practice. In our adaption of the HJ framework based on Almeida and Schneider (2021), the model identified through the pricing equations is the sum of a linear part, that is identical to the one from HJ, and a nonlinear part, featuring minimal volatility and no price impact. This sum is positive in every state of the world, and represents a valid stochastic discount factor. Economically, the nonlinear part can be interpreted as the minimal-cost nonlinear trading strategy (insurance policy) that a representative agent optimizing her Sharpe ratio portfolio would need to enter in order to price all assets in question with a linear model, as well as to ensure the absence of arbitrage.

To work with RMP empirically, we estimate conditional expectations from realized returns and conditioning information through novel machine learning techniques, imposing economically motivated restrictions in the form of absence of arbitrage opportunities. While we can infer RMP also unconditionally, conditional RMP enables us to better investigate its relation to market events and conditional risk premia. We show that insuring against mispricing becomes more expensive during market turmoil. This is consistent with the notion that linear asset pricing models perform well on
average and during normal times, whereas they imply larger (residual) mispricing during crisis periods, when the marginal utility of investors is high.

We exploit the weak data requirements of RMP to investigate the cross section of aggregate equity and short-term bonds in multiple markets linked by their respective exchange rates. Our motivation for this is twofold. First, Verdelhan (2018) documents that the strong correlation structure of bilateral exchange rates can be attributed to essentially two risk factors, carry and dollar, accounting for a large portion of the share of systematic variation. Second, Koijen and Yogo (2020) postulate that when determining their optimal demand, global investors hold exchange rate exposure not only through bonds, but also equity. Hence, with U.S. as our base currency, we compute RMP for United Kingdom, Switzerland, Australia, Canada, Japan, New Zealand, and the Euro area, and additionally examine currency returns with respect to the U.S. dollar.

We summarize our empirical findings as follows. First, we document a strong link between RMP and conditional expected returns. Specifically, we find that the cross section of domestic and international aggregate equity returns, as well as currency returns, are explained by RMP even after controlling for conventional linear factors, both in-sample and out-of-sample. Importantly, we show that the price of risk is positive for conditional expected returns, whereas for unconditional (realized) returns, we obtain the typical negative relation between risk and return previously documented in the literature (see, e.g., Ang, Hodrick, Xing, and Zhang (2006), among others). We conclude that studying conditional expected returns is instrumental for inference and estimation.

Second, we investigate the potential determinants of RMP. We find that RMP is related to extant measures of uncertainty, such as financial uncertainty of Jurado, Ludvigson, and Ng (2015), and the VIX. Next, we show that RMP is related to nonlinear proxies of intermediary leverage constraints and exhibits a close link with market-wide funding liquidity shocks, as measured by the TED spread. Still, in regression analysis, the explanatory power varies from 4% to 28% in both domestic and international markets, suggesting that RMP can be interpreted as a non-redundant measure of financial distress. Our results hold also out-of-sample, implying that the evidence provided is not likely to be spurious or driven by overfitting in sample. Overall, we document that residual mispricing is high during periods of uncertainty and during episodes of low (funding) liquidity, or financial distress. Our findings thus confirm economic intuition that nonlinearities in asset pricing models ensuring absence of arbitrage opportunities are particularly relevant during financial turmoil.

Third, we investigate whether the RMP index is related to measures of observed statistical arbitrage violations or mispricing in the data. To this end, we provide
a comparison with the market dislocation index (MDI) of Pasquariello (2014), that captures arbitrage violations in international markets. In addition, following Du, Tepper, and Verdelhan (2018b) we consider covered interest rate parity (CIP) violations observed for the developed markets in our sample. We find that RMP, constructed as an ex-ante measure consistent with no-arbitrage, explains and predicts financial market dislocations and CIP deviations at different tenors, varying from three months to ten years. Importantly, our findings hold also out-of-sample, confirming the robustness of the construction and estimation of RMP, and its applicability as an economic indicator. Hence, we conclude that RMP contains additional information not subsumed by proxies of intermediaries’ constraints or existing deviations from the no-arbitrage condition, such as CIP violations.

RMP lends itself also to assess the predictions of parametric economic models. Studying the behavior of (unconditional) RMP through the lens of the Colacito and Croce (2013) model, we find that RMP is high and positive during distress periods, i.e. whenever both domestic and foreign returns are negative. This confirms both the intuition, as well as the nonparametric evidence gathered from our data. Hence, we illustrate that nonparametric RMP shares similar traits of this parametric benchmark model incorporating long run risk components, at least unconditionally.

Overall, we document that RMP is high during periods of uncertainty and during episodes of low (funding) liquidity, exhibiting a substantial effect within and across different financial markets. We show that augmenting the set of assets available to investors to account for nonlinear payoffs decreases the magnitude of the underlying residual mispricing. Our findings thus contribute to the study of the relation between uncertainty and mispricing, by arguing that nonlinearities are important for the absence of arbitrage opportunities, regardless of the assets being priced. Naturally, these nonlinearities are likely more important for assets exhibiting crash risk (such as currencies).

Literature. Our framework has implications for, and builds on, several relevant strands in the literature surveyed by Brunnermeier, Farhi, Kojjen, Krishnamurthy, Ludwigson, Lustig, Nagel, and Piazzesi (2021): international finance, expectation formation, and intermediary asset pricing. Methodologically, it follows the call in Nagel (2021) to adapt machine learning techniques with structural economic constraints.

Our first building block is the literature examining unspanned risks, that is, factors that affect risk premia, but do not appear in prices. Collin-Dufresne and Goldstein (2002) and Duffee (2011) have suggested these unspanned risks to be important to reconcile prices with risk premia. Further such studies include Gabaix (2012), Joslin, Priebsch, and Singleton (2014), and Filipović, Larsson, and Trolle (2017). The
advantage of our approach stems from the formulation of a minimum-variance SDF directly in terms of both a linear market model, and a portion that is not observed in prices, that we show to be related to conditional risk premia.

We select the stochastic discount factor (SDF) according to a minimum dispersion criterion (Hansen and Jagannathan, 1991; Almeida and Garcia, 2012, 2017), in a particular polynomial space proposed by Almeida and Schneider (2021), that allows us to control precisely the nonlinearity needed to ensure positivity of the SDF (similar techniques are used in Lasserre, 2010; Renner and Schmedders, 2015). Differently from the positive Hansen and Jagannathan (1991) SDF that is linear in option-payoff-type functions, our model is formulated in terms of a part that is linear in returns, and another part of minimum divergence that is at least quadratic. Sandulescu, Trojani, and Vedolin (2021) employ unconditional minimum dispersion SDFs in an international setting and study the relationships between market incompleteness, financial market structures and international finance puzzles. We depart from this framework by constructing conditional minimum-variance SDFs that ensure absence of arbitrage in every state of the world. Moreover, we document the time series properties of conditional risk premia and the importance of time-varying weights in optimal portfolios, while focusing on the residual mispricing relative to linear factor models.

Our paper is also related to a strand of literature that studies the cross-section of currency returns predictability using different risk factors, such as global exchange rate risk (Lustig, Roussanov, and Verdelhan (2011), Verdelhan (2018), Colacito, Croce, Gavazzoni, and Ready (2018)), global imbalances (Corte, Riddiough, and Sarno (2016)), macro fundamentals (Colacito and Croce (2011), Hassan (2013), Gabaix and Maggiori (2015), Colacito, Riddiough, and Sarno (2020)), downside risk (Lettau, Maggiori, and Weber (2014)), among others. We contribute to these studies by extending the benchmark linear models to accommodate a nonlinear factor that renders international markets arbitrage-free. We show that our nonlinear factor is non-redundant and explains the cross-section of currency returns, also out-of-sample. Moreover, we focus on conditional expected returns.

Lastly, we obtain conditional risk premia with novel machine learning techniques based on distribution embedding in reproducing kernel Hilbert spaces (RKHS), first introduced in Song, Huang, Smola, and Fukumizu (2009). Using reproducing kernel Hilbert spaces (RKHS), we are able to incorporate the information from multiple conditioning variables simultaneously for conditional expectations of a multitude of functions, allowing us to discern between the role of the information set, and
the information induced by no-arbitrage.\footnote{RKHS are particularly relevant for statistical learning, or finding a predictive function based on data. They can be thought of as a generalization of linear regressions, in that they reveal optimal functions that represent data, rather than optimal coefficients, while maintaining the computational efficiency of OLS. Differently from other machine learning techniques such as neural nets, RKHS are a white, rather than a black box, in that all steps and expressions are deterministic and interpretable, and the researcher operates always at unique, and globally optimal estimates.} In contrast to linear regressions, such as ordinary least squares, we do not have to make assumptions regarding the functional form, as the optimal function itself is found to accurately represent the data. In particular, we can accommodate nonlinear interactions between returns and conditioning variables, whereas if the true relation between them was linear, RKHS would correctly approximate it. Relative to the local polynomial regressions in Nagel and Singleton (2011) and Roussanov (2014), RKHS distribution embedding is a machine learning technique that performs well in small samples. Furthermore, as opposed to neural nets, it always operates at globally optimal estimates, with an exactly known functional form of the resulting conditional expectations. Lastly, by estimating conditional expectations, we avoid the noise induced by rolling windows. We apply this technique in particular to obtain out-of-sample inference. To the best of our knowledge, the only two prior studies to use RKHS in financial economics are Kozak (2020), to estimate coefficients of a linear pricing kernel, and Schneider (2021) to estimate discrete probabilities on low-dimensional scenario trees. In our empirical analysis, we account for the interest rate differential and realized volatility as conditioning variables, conforming with Lustig, Roussanov, and Verdelhan (2011) and Menkhoff, Sarno, Schmeling, and Schrimpf (2012), among others.

The rest of the paper is organized as follows. Section 2 provides an introductory example. Section 3 describes the framework employed to assess residual mispricing. Section 4 presents the data and our empirical findings. Section 5 contains four robustness checks to our main empirical analysis. In Section 5.1 we study RMP when nonlinear claims are added to the linear return model. Section 5.2 compares RMP with the mispricing factors from Stambaugh and Yuan (2017), and in Section 5.3 we perform additional robustness results for the relation of RMP with expected returns. In Section 5.4, we compute RMP under a long-run-risk model, and investigate its behavior under two extreme conditioning events. Section 6 concludes. Appendices A and B contain details of the construction of RMP and the estimation of conditional expectations. Lastly, Appendices C and D include additional results and an illustration of the time-series properties of time-varying weights, respectively.
2 Introductory example

To develop intuition for the arguments to come, imagine an economy accommodating two assets with risky excess returns $r_X$ (the market) and $r_Y$, with standard deviations $\sigma_X$ and $\sigma_Y$, respectively, and a riskless bond. Starting from a linear CAPM-type model for the SDF, $M^* := a + b r_X$, we obtain the well-established relation,

$$
\mathbb{E} [r_Y] = \frac{\text{Cov}(M^*, r_Y)}{\text{Cov}(M^*, r_X)} \mathbb{E} [r_X] \quad (1)
$$

between the two risky excess returns. Note that $M^*$ is a portfolio consisting of a riskless bond with payoff $a$, and of $b$ times the market return $r_X$. The popularity of the linear pricing approach likely stems from the observation that the expected return prediction of the linear factor model agrees with the prediction from ordinary least squares (OLS), with the additional assertion that the intercept is zero.

The elegance of the unification of pricing and OLS regression comes with a drawback, however, since $M^*$ is in general not a SDF. As a linear function of $r_X$, that as an excess return takes both positive and negative values, it becomes negative with positive probability. Intuitively, $M^*$ guarantees that the law of one price is satisfied, however there exist potential arbitrages to be explored.

To remedy this problem, consider adding a nonlinear payoff $M^0$ to $M^*$, to specify a valid SDF $M$ as

$$
M := M^* + M^0.
$$

To be consistent with pricing in this economy, $M^0$ must satisfy a number of basic restrictions. First, it ought to be mean-zero, $\mathbb{E} [M^0] = 0$, so that the price of the risk-free bond is unaltered, and likewise orthogonal to $r_X$, i.e. $\mathbb{E} [r_X M^0] = 0$, to not alter the pricing of $r_X$. Second, to leave pricing relation (1) unchanged, it needs to satisfy $\mathbb{E} [r_Y M^0] = 0$. Third, and most importantly for the context of this paper, it needs to be such that $M^* + M^0 \geq 0$ in every state of the world. This constraint guarantees the absence of arbitrage opportunities from the fundamental theorem of asset pricing.

We can price the claim $M^0$ with the SDF $M$, such that $\mathbb{E} [(M^* + M^0) M^0] = \mathbb{E} [M^0^2] \geq 0$, with guaranteed positive payoff when $M^*$ becomes negative, since $M^0 \geq -M^*$. Note that the expected profit, or risk premium, from this trading strategy is

$$
\mathbb{E} [M^0] - \mathbb{E} [M^0^2] = -\mathbb{E} [M^0^2] \leq 0. \quad (2)
$$
Together with the property that it yields a positive payoff if \(M^*\) is negative, it represents an insurance strategy, insulating the linear pricing model \(M^*\) against arbitrage, with corresponding insurance premium \(\mathbb{E}[M^{o2}]\).

To sustain the linear structure of \(M^*\) (in the coefficients), the most natural way of satisfying all the requirements listed above reads

\[
M^o = w_0 + w_1 r_X + w_2 r_X^2. \tag{3}
\]

In this way, simple conditions on the weights \(a, b, w_0, w_1, w_2\) can be imposed to ensure positivity of \(M^o\). If the discriminant \((b + w_1)^2 - 4w_2(a + w_0)\) of the quadratic polynomial in \(r_X: M = M^* + M^o = (a + w_0) + (b + w_1)r_X + w_2 r_X^2\) is non-positive, then \(M \geq 0\) in every state of the world, and thus represents a valid SDF. From the aforementioned constraints, the weights \(w_0, w_1, w_2\) must additionally solve

\[
\begin{align*}
\mathbb{E}[w_0 + w_1 r_X + w_2 r_X^2] &= 0, \\
\mathbb{E}[r_X(w_0 + w_1 r_X + w_2 r_X^2)] &= 0, \text{ and} \\
\mathbb{E}[r_Y(w_0 + w_1 r_X + w_2 r_X^2)] &= 0.
\end{align*}
\]

If the discriminant is positive, a higher even-order polynomial in (3) can be chosen to satisfy the linear constraints and to ensure positivity. In any case, a unique and conservative set of coefficients can be chosen by minimizing the second moment of \(M^o\) subject to the constraints above, yielding the cheapest strategy insulating the linear model against arbitrage violations.

To see the contradictions and arbitrage opportunities that can arise by taking \(M^*\) as a SDF without insurance, suppose that \(b\) is negative, \(a\) is positive, and consider a call option on \(r_X\) with strike price \(K := a/|b|\), which is precisely the cutoff from which \(M^*\) becomes negative if crossed by \(r_X\).\(^2\) The payoff from the call option is \((r_X - K)^+\), with \((z)^+\) denoting the positive part of \(z\), and the price\(^3\) using only \(M^*\) reads

\[
\mathbb{E}[M^*(r_X - K)^+] = \mathbb{E}\left[\left(a + br_X\right)\left(r_X - \frac{a}{|b|}\right)^+\right] \leq 0, \tag{4}
\]

since the payoff of the call option is positive, precisely when \(M^*\) is negative. Taking together the non-negative payoff of the call option with the non-positive price computed with \(M^*\) constitutes a (weak) arbitrage portfolio.

\(^2\)This can be seen from \(a + br_X \leq 0 \iff r_X \geq a/|b|\) for \(a \geq 0, b < 0\). Different signs of \(a\) and \(b\) would lead to different arbitrage portfolios. For instance, with \(b > 0\), we could work with a put option.

\(^3\)With options written on \(r_X\), the payoff could even be replicated from option portfolios according to Carr and Madan (2001). Otherwise, it would be the cheapest theoretical insurance strategy insulating the linear model against arbitrage, engineered from existing payoffs.
On the other hand, the price computed with \( M = M^* + M^\circ \) reads

\[
\mathbb{E} \left[ M \left( r_X - \frac{a}{|b|} \right)^+ \right] = \mathbb{E} \left[ (M^* + M^\circ) \left( r_X - \frac{a}{|b|} \right)^+ \right] \geq 0,
\]

since \( M \) is non-negative in every state of the world, and so is \( \left( r_X - \frac{a}{|b|} \right)^+ \), ruling out arbitrage.

The modification \( M^\circ \) is the most conservative (since it has minimal variance) residuum that makes \( M = M^* + M^\circ \) a valid SDF without changing the linear pricing equations. As such, it is interpreted as residual mispricing, in the sense that it measures the minimal distance between the linear pricing model \( M^* \) and the closest no-arbitrage model \( M \).

3 Conditional residual mispricing

In this section, we describe our modeling assumptions to assess the degree of residual mispricing implied by different asset markets in detail. Following our introductory example in the previous section, we build our approach upon a unifying framework between standard linear asset pricing models, and the notion of no-arbitrage. No-arbitrage implies the existence of a strictly positive stochastic discount factor (SDF), but SDF models that are linear in excess returns can not be positive in every state of the world. Below, we develop the discrepancy between the linear model and the closest positive SDF as our RMP index, and motivate why it measures mispricing.

To this end, we reformulate the Hansen and Jagannathan (1991) asset pricing framework, such that linear SDF models are merely projections onto the linear asset span and thus naturally accommodate linear asset pricing models, while being positive in all states of the world. We achieve this task by working with a family of SDFs that are non-negative polynomials in the underlying asset excess returns. We are not the first to consider polynomial pricing kernels, as starting from Kraus and Litzenberger (1976), a large literature is built on power series expansions (see, e.g. Harvey and Siddique, 2000; Dittmar, 2002, among others). However, our approach is the first to yield positive polynomial SDFs, to the best of our knowledge. This is not an arbitrary choice, as we are interested in insuring linear models against negativity, and this is naturally done with polynomials.\(^4\) Moreover, this specification allows for an additive decomposition: the linear factor model plus the nonlinear part. This decomposition is appealing because the nonlinear part is constructed to be orthogonal to the factor

\(^4\)From an econometric point of view, under mild technical conditions, polynomials are a basis of the space of square-integrable random variables under the real-world probability \( \mathbb{P} \) and thus also well-suited for approximation.
model, allowing us to isolate the differential effect. As the polynomial degree becomes large, the specification converges to the true, unknown SDF. While they are positive by construction, exponentially affine models share none of these appealing properties.

In incomplete arbitrage-free markets, there may be many candidate SDFs. Following the literature, we therefore choose a representative of this set that exhibits minimal variance. It is well known that such a minimum-variance SDF is also associated with the sharpest upper bound for the Sharpe ratio of all trading strategies in this economy (see Hansen and Jagannathan, 1991).

With $R_{0,t+1}^i$ the one-period riskless gross bond return in economy $i$ known at time $t$, and $R_{1,t+1}^i, \ldots, R_{m_i,t+1}^i$ gross returns from time $t$ to $t + 1$, we select the SDF pricing excess returns $R_{i,t+1} := (R_{1,t+1}^i - R_{0,t+1}^i, \ldots, R_{m_i,t+1}^i - R_{0,t+1}^i)^\top$ according to the following optimization problem,

$$\begin{align*}
\text{minimize} & \quad \mathbb{E}_\mathbb{P} \left[ M_{t+1}^2 \middle| \mathcal{F}_t \right], \\
\text{subject to} & \quad \mathbb{E}_\mathbb{P} \left[ M_{t+1} R_{i,t+1} \middle| \mathcal{F}_t \right] = 0_{m_i}, \quad \mathbb{E}_\mathbb{P} \left[ M_{t+1} \middle| \mathcal{F}_t \right] = \frac{1}{R_{0,t+1}^i}, \\
M_{t+1} & \geq 0, \quad \text{P.a.s.}
\end{align*}$$

(5)

where $0_{m_i}$ denotes a zero vector of length $m_i$, $\mathcal{F}_t$ is the information set available to investors at time $t$, and $P_{n,m_i}$ denotes the set of polynomials in $R_{i,t+1}$ with fixed even degree $n$. We denote the optimal solution$^5$ in Equation (5) by $M_{i,t+1}$, with $i \in \{d, f\}$. It is linear in powers of $R_{i,t+1}$, and denoting the $j$-th entry of $R_{i,t+1}$ by $r_{i,j}$, and $B_{i,n,t+1} := (1, r_{i,1}, \ldots, r_{i,m_i}, r_{i,1}^2, r_{i,1} r_{i,2}, \ldots, r_{i,m_i}^m)^\top$, we can write it as

$$M_{i,t+1} = w_{i,t}^\top B_{i,n,t+1},$$

(6)

where the optimal weights $w_{i,t}$ are known at time $t$. Note that the optimization problem (5) ensures the absence of arbitrage in every state of the world. This requirement is stricter than the absence of statistical arbitrage in the sense of the Ross (1976) arbitrage pricing theory, and Chamberlain and Rothschild (1983) that is used in studies with larger cross sections (for instance Haddad et al., 2020; Kozak et al., 2020; Chen et al., 2021).

The original construction in Hansen and Jagannathan (1991) performs the optimization in $L^2_\mathbb{P}$, the space of square-integrable random variables under the real-world probability measure $\mathbb{P}$. We perform the optimization in formulation (5) on a

---

$^5$This optimization problem is considered in Almeida and Schneider (2021) to engineer positive polynomial likelihood ratios. Since problem (5) is convex, it can be solved rapidly with a unique optimal solution. We describe the corresponding finite-dimensional program in detail in Appendix A.
smaller set than in the original formulation. The reason for this simplification is twofold. Firstly, we can accommodate linear models in this way. Linear models, such as the one implied by the CAPM, or the Fama French factor models are widely used in industry and academia. Secondly, we can easily control the condition that the SDFs are positive, through a simple sum of squares condition that is available for positive polynomials. Positive Hansen and Jagannathan (1991) SDFs are linear in option-type functions of the returns, hard to estimate conditionally, and not easy to reconcile with practitioners’ models. Moreover, we can uniquely decompose the solution $M_i$ as

\begin{align}
M_{i,t+1} &= M_{i,t+1}^* + M_{i,t+1}^o = \underbrace{\mathbf{w}_{i,t}^\top \mathbf{B}_{i,n,t+1}}_{\text{HJ linear part}} + \underbrace{\mathbf{w}_{i,t}^o \mathbf{B}_{i,n,t+1}}_{\text{arbitrage insurance payoff}}, \quad (7)
\end{align}

where $M_{i,t+1}^*$ is linear in $\mathbf{R}_{i,t+1}$, and identical to linear Hansen-Jagannathan SDFs with no positivity constraint. Naturally, $\mathbf{w}_{i,t}^*$ has zero entries on all nonlinear components of $\mathbf{B}_{i,n,t+1}$. The second component $M_{i,t+1}^o$ is the random variable with smallest variance, such that $M_{i,t+1}$ is positive. Furthermore, it is orthogonal to $M_{i,t+1}^*$ and all its components. As a second benefit of orthogonality, the second moment of $M_{i,t+1}$, an upper bound for the conditional Hansen-Jagannathan bound, can be decomposed into

\begin{align}
\mathbb{E}^P [M_{i,t+1}^2 | \mathcal{F}_t] &= \mathbb{E}^P [M_{i,t+1}^*^2 | \mathcal{F}_t] + \mathbb{E}^P [M_{i,t+1}^o^2 | \mathcal{F}_t], \quad (8)
\end{align}

where $\mathbb{E}^P [M_{i,t+1}^2 | \mathcal{F}_t]$ can be interpreted as the marginal Sharpe ratio increase arising from nonlinear claims necessary for hedging arbitrage portfolios. Accordingly, we define the insurance profit as

\begin{align}
RM P_{i,t+1} := M_{i,t+1}^o - \mathbb{E}^P [M_{i,t+1}^o^2 | \mathcal{F}_t], \quad (10)
\end{align}

as the object of interest of our paper. The trading profit $RM P_{i,t+1}$ is the most conservative measure of mispricing relative to the linear SDF model, since $M_{i,t+1}^*$ is the modification to the linear SDF $M_{i,t+1}^*$ with the smallest variance ensuring no-arbitrage.
It is a measure of residual mispricing, since it reflects the portion of the positive SDF that is missed by the linear model to guarantee the absence of arbitrage. As motivated in the introductory Section 2, by construction

$$
\mathbb{E}^P \left[ RMP_{t,t+1} \mid \mathcal{F}_t \right] = -\mathbb{E}^P \left[ M_{t,t+1}^{\circ 2} \mid \mathcal{F}_t \right] \leq 0.
$$

(11)

The presence of conditional expectations in formulation (5) necessitates a model. Rather than working with a parametric formulation, or nonparametric local linear or local constant regressions, we learn the conditional expectation function in reproducing kernel Hilbert spaces (RKHS). RKHS are a powerful framework used in machine learning to transform potentially infinite-dimensional features of an object to be studied through data, such as a conditional expectation, to simple finite-dimensional linear algebra operations through the so-called kernel trick. In financial economics they are to this date rarely used (a notable exception is Kozak, 2020, who uses RKHS to learn the coefficients of a linear SDF). Learning the conditional expectation function requires additional steps over the standard RKHS machinery. Importantly, one must specify which variables make up investors’ information set $\mathcal{F}_t$. Contrary to local constant and local linear nonparametric regressions, RKHS conditional distribution embedding allows us to use multidimensional conditioning information with comparably short time series. In our empirical application, we are able to use a bivariate information set, despite a relatively small number of data points ($\approx 540$).

4 RMP in international markets

In this section, we describe the data employed in the empirical analysis, the estimation procedure, and study the time series properties of international RMP indices. We then explore potential determinants of RMP, and show that RMP is related to arbitrage violations in the data in both contemporaneous and predictive regressions. Subsequently, we document a sizeable positive premium for RMP in the cross-section of international stocks, as well as for currencies. Finally, we analyze the link between RMP and exchange rates.

4.1 Data

We obtain price data for MSCI country indices, exchange rates and one-month deposit rates from Datastream, for the period spanning January 1985 to June 2020. Due to

\[ 10 \]

We describe the procedure along with its in-sample and out-of-sample cross validation in Appendix B.

\[ 11 \]

We extend the sample period to January 1975 in our out-of-sample exercise.
RMP’s minimal data requirements, we are able to examine eight developed economies, where we consider the U.S. as domestic market, and we define as foreign markets: the United Kingdom, Switzerland, the Euro area\textsuperscript{12}, Japan, Canada, Australia and New Zealand. Lastly, we derive aggregate equity returns at the country level and risk-free rates from monthly deposit rates, and focus on bilateral trading.

With international data in mind, we follow Lustig, Roussanov, and Verdelhan (2011) and Menkhoff, Sarno, Schmeling, and Schrimpf (2012), and use the interest rate differential, as well as exchange rate volatility as conditioning variables. We calculate two sets of RMP indices, one using the information of the entire sample (in-sample), and the other one using only information available prior to each data point (out-of-sample).

4.2 Estimation procedure and market structure

In our estimation, we assume that investors can trade both equities and short-term bonds internationally, through the exchange rate $X$. Markets are thus integrated internationally. We follow extant literature and study linear factor models in bilateral exchange rates (see, e.g. Backus, Foresi, and Telmer (2001), Lustig, Roussanov, and Verdelhan (2014) and Verdelhan (2018), among others). The results that we derive are hence based on bilateral trading, assuming that U.S. is the domestic economy. Consequently, the conditional SDFs that we derive correctly price domestic and international equities and short-term bonds. Specifically, for domestic U.S. investors, the vector of tradable returns reads $\mathbf{R}_{d,t+1} := (R_{d1,t+1} - R_{d0,t+1}, R_{d2,t+1} - R_{d0,t+1}, R_{d3,t+1} - R_{d0,t+1})^\top$, as well as the one-period domestic riskless gross bond return $R_{d0,t+1}$, where $R_{d1,t+1}$ is the domestic market gross return, and $R_{d2,t+1} = R_{d0,t+1}X_{t+1}$ is the foreign risk-free rate in domestic units, and $R_{d3,t+1} = R_{d0,t+1}X_{t+1}$ is the foreign market gross return in domestic currency. Similarly, for foreign investors, the vector of tradable returns is $\mathbf{R}_{f,t+1} := (R_{f1,t+1} - R_{f0,t+1}, R_{f2,t+1} - R_{f0,t+1}, R_{f3,t+1} - R_{f0,t+1})^\top$, as well as the one-period foreign riskless gross bond return $R'_{f0,t+1}$, where $R'_{f1,t+1}$ is the foreign market gross return, and $R'_{f2,t+1} = R'_{f0,t+1}/X_{t+1}$ is the foreign market gross return in foreign currency.\textsuperscript{13}

In order to estimate conditional expectations, we employ RKHS embedding of conditional distributions. Estimating functions such as conditional expectations in RKHS is a machine learning technique that is particularly tractable, and consistent in an econometric sense (Boudabsa and Filipović, 2020). Through the so-called kernel

\textsuperscript{12}We use Germany as a proxy for the Euro area prior to the introduction of the euro currency.

\textsuperscript{13}Note that we are not pricing the foreign exchange forward, for which recent studies (see, e.g. Du, Tepper, and Verdelhan (2018b)) have found deviations from covered interest parity, nor do we consider interbank or money market rates, for which these deviations are usually documented (Rime, Schrimpf, and Syrstad (2022)), but only government bonds and spot exchange rates.
trick, infinitely-dimensional features can be processed through the evaluation of a carefully chosen inner product, leading to simple solutions in terms of numerical linear algebra operations. Analogously to ordinary least squares, that recovers optimal coefficients that represent the data in a linear model, RKHS help recover optimal functions that best describe the data. Intuitively, if the correct representation was linear, then RKHS would closely approximate it. With various constructions of conditional expectations in RKHS in the literature (Song et al., 2009; Grünewälder et al., 2012; Park and Muandet, 2020), their computation necessitates a parameter for the kernel of the so-called input space, the conditioning variables, as well as a regularization parameter. The kernel parameter controls the smoothness of the family of functions represented, while the regularization parameter controls the trade-off between smoothness and the error of the conditional expectation fitted to the realizations. In our estimation, we implement a Gaussian kernel featuring one (volatility) parameter.\textsuperscript{14} Similarly to local linear and local constant nonparametric regressions, we tune the parameters in a cross validation procedure as described in Appendix B. Importantly, the two aforementioned parameters are used for conditional expectations of any function that is an element of the RKHS, making the procedure very parsimonious with respect to the number of parameters. Since the RKHS representation of conditional expectations follows a sequence of simple linear algebra operations, it is also a very fast procedure. For the in-sample RMP, we perform time series cross-validation over the entire sample, while we repeat the out-of-sample cross validation for each data point (using only information prior to the data point). We also consider our program (5) in terms of unconditional expectations estimated through sample averages, and refer to the corresponding residual mispricing index as the unconditional RMP.

With (un)conditional expectations at hand, we solve the minimization problem (5) for each data point as a conic program, detailed in Appendix A, for a unique optimal polynomial \( M^\ast_{i,t+1} + M^0_{i,t+1} \).\textsuperscript{15} The linear \( M^\ast_{i,t+1} \) is obtained as the classical Hansen and Jagannathan (1991) solution without positivity. The time series of the RMP index is then obtained by evaluating the polynomial \( M^0_{i,t+1} \) at the return data \( R_{i,t+1} \) realized at time \( t+1 \), and weights estimated using the information set at time \( t \), and subtracting its price, as in (10), reusing the conditional expectations used in the initial optimization. In Section 4.3, we describe how the RMP index evolves over time, and how it differs between the conditional and the unconditional formulation.

Owing to our international sample comprised of seven currencies GBP, CHF, JPY,  
\textsuperscript{14}This choice of kernel is standard in the literature. Our results are not driven by a particular kernel function.  
\textsuperscript{15}This optimization is performed in a fraction of a second with a conic solver such as MOSEK.
EUR, AUD, CAD, and NZD, we distinguish between domestic (U.S.), and international RMP. Since our estimation is carried out for bilateral trading, we take domestic (U.S.) RMP as the average

\[ RMP_{dom,t+1} = \frac{1}{7} (RMP_{USDGBP,t+1} + \cdots + RMP_{USDNZD,t+1}), \]  

(12)

and international RMP computed as\(^{16}\)

\[ RMP_{int,t+1} = \frac{1}{7} (RMP_{GBPUSD,t+1} + \cdots + RMP_{NZDUSD,t+1}), \]  

(13)

in what follows, and analogously for \(\pi\), the trading profit from the linear model defined in Equation (9).

### 4.3 Time series properties of RMP

In this section, we describe the statistical properties of the RMP index. Figure 1, top panel, shows the time series of realized U.S. domestic \(RMP_{dom}\) from (12). As constructed, RMP has an (un)conditional mean slightly below zero, reflecting the insurance premium demonstrated in (11). Realizations can be both positive and negative to complement the linear \(M^*_dom\), such that absence of arbitrage is ensured in all states of the world. On average, and whenever RMP is close to zero, the linear model provides a good approximation of the return data, such that the insurance profit is small. It is noteworthy that the RMP realizations with the largest magnitudes occur during crisis times and are uniformly positive, in line with the insurance interpretation in our introductory example in Section 2, or as a financial distress measure. The largest realization occurred during the 1987 crisis, while the Black Wednesday, the Asian and LTCM crises feature comparable values. RMP during the 2008 financial crisis is as large as during the recent COVID-19 crisis, suggesting that we are experiencing new record highs since the strikingly low residual mispricing risk over the past decade. We plot international RMP (13) in the bottom panel of Figure 1, computed as the average across all foreign countries (excluding the U.S.), and document patterns similar to those of U.S. RMP. Importantly, this high similarity in the time series of the two indices (featuring a correlation of 92%) suggests that U.S. residual mispricing provides a good approximation of aggregate international residual mispricing (and vice versa).\(^{17}\)

\(^{16}\)Transforming first RMP in U.S. dollars and then taking the average yields virtually identical results. The two indices feature perfect positive correlation.

\(^{17}\)On the other hand, RMP is mildly persistent, and a question arises whether this is due to the conditional expectations used in the computation of the RMP. We rule out this possibility, as Figure C1 shows in a scatter plot that the realizations of the unconditional vs. the conditional SDFs do not deviate systematically.
Indeed, the co-movement of RMP indices between domestic and individual foreign markets is high as well (see, e.g. Appendix Table C1).

Figure 1: Time series of international residual mispricing

![Time series of international residual mispricing](image)

This figure plots the time series of the domestic U.S. RMP (top panel) and the international RMP, computed as the average across all foreign countries excluding U.S. (bottom panel). Gray bars denote NBER recessions. The sample period is January 1985 to June 2020.

4.4 Determinants of RMP

In this section, we explore potential determinants of residual mispricing, both domestically and internationally. Guided by the countercyclical patterns documented for RMP, we start by studying its relation with financial uncertainty proxies in international markets. Uncertainty data containing monthly proxies for the macro economy and the financial sector are obtained from Jurado, Ludvigson, and Ng (2015) and Ludvigson, Ma, and Ng (2015). Specifically, they extract an economic uncertainty index as the conditional volatility of the unpredictable component of a large number of economic indicators. To further examine the link between RMP and uncertainty in international markets, we use the VIX, a volatility index provided by CBOE, as an alternative measure of uncertainty.\(^{18}\)

Figure 2 (top panel) plots the time-series of U.S. RMP and proxies of financial uncertainty, illustrating a strong positive relation, especially during periods of financial

\(^{18}\)For the sample period between January 1985 and December 1989 we use monthly data extended for S&P 500 implied volatility from Berger, Dew-Becker, and Giglio (2020).
distress. To explore their link more systematically, we regress RMP on financial uncertainty proxies. Specifically, our empirical specification reads

$$RMP_{i,t} = \alpha_i + \beta_i X_t + \epsilon_{i,t},$$

where $X$ is a proxy for financial uncertainty, and $i \in \{\text{dom, int}\}$.

**Figure 2: Determinants of RMP**

This figure plots the time series of the U.S. RMP index (dashed-line) against (i) the financial uncertainty index of Jurado, Ludvigson, and Ng (2015) (Top left Panel); (ii) VIX (Top right Panel); (iii) the intermediary leverage ratio squared, or the squared reciprocal of the capital ratio of the intermediary sector from He, Kelly, and Manela (2017) (Bottom left Panel); (iv) TED spread (Bottom right Panel). Gray bars denote NBER recessions. The sample period is January 1985 to June 2020 (January 1986 to June 2020 for TED spread).

Table 1 reports the estimation results. All betas are positive and statistically significant at the 1% level, suggesting that periods when the underlying financial uncertainty is high are accompanied by larger residual mispricing. The explanatory power is similar for U.S. RMP and international RMP (which is the average across international markets excluding the U.S.), with $R^2$ of 19.2% and 19.6%, respectively. Similarly, we find a strong positive relation between RMP and the VIX. On average, VIX accounts for 24.9% (25.4%) of the variation of RMP in domestic (international) economies. Both, the financial uncertainty index, and VIX, remain significant when

\[\text{For ease of interpretation and comparability of coefficients, we standardize all variables to have mean zero and unit variance.}\]
running multivariate regressions. Overall, we document a close link between residual mispricing as captured by the RMP index and financial uncertainty in international markets. Thus, RMP can be interpreted not only as a measure consistent with the absence of arbitrage opportunities, but also as a financial distress measure. Since existing proxies only capture about 20-25% of the variation in RMP, the latter is non-redundant, while being economically grounded and theoretically motivated by the no-arbitrage condition. Moreover, since the RMP index uses only return data on international equities and bonds, it is readily available for several economies, without having to rely on options data, for instance, or having to aggregate a considerable number of economic indicators (and estimation errors) to derive a measure of uncertainty.

Table 1: Determinants of RMP

This table reports the estimated coefficients from regressing the RMP index on various determinants, in the i = domestic (12), and international market (13): \( RMP_{i,t} = \alpha_i + \beta_i X_t + \epsilon_{i,t} \). Proxies for financial uncertainty are the financial uncertainty index of Jurado, Ludvigson, and Ng (2015) and the VIX; intermediary leverage ratio squared is from He, Kelly, and Manela (2017), whereas TED spread is from FRED. The domestic market is the U.S., whereas international is the average across foreign markets excluding the U.S. All variables are standardized. Data are monthly and span January 1985 to June 2020 (January 1986 to June 2020 for the TED spread). Newey-West standard errors are reported in brackets. *** and ** indicate significant at the 1% and 5% level, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Panel A: U.S. Residual Mispricing</th>
<th>Panel B: International Residual Mispricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial uncertainty</td>
<td>0.441***</td>
<td>0.392***</td>
</tr>
<tr>
<td>VIX</td>
<td>0.560***</td>
<td>0.490***</td>
</tr>
<tr>
<td>Intermediary leverage</td>
<td>0.234***</td>
<td>-0.006</td>
</tr>
<tr>
<td>TED spread</td>
<td>0.258**</td>
<td>0.132</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>19.2% 24.9% 5.3% 6.5% 21.8% 26.8%</td>
<td>19.6% 25.4% 4.6% 8.2% 22.6% 27.7%</td>
</tr>
</tbody>
</table>

We proceed by exploring additional potential sources of residual mispricing. To document the ability to trade assets with nonlinear payoffs, that are consistent with RMP, we investigate the link with financial intermediaries, and in particular periods when their ability to provide liquidity is hindered. The reason we focus on financial intermediaries is twofold. First, they represent specialized investors, who are in the unique position of trading most asset classes, especially the more complex ones, that trade over-the-counter (OTC). Second, we can only measure the propensity to trade of intermediaries, and periods when they experience financial constraints such that their risk bearing capacity is impaired, are likely to coincide with periods of heightened degrees of unpriced risk, potentially leading to limits to arbitrage and subsequent dislocations in financial markets.

---

20Since financial uncertainty index and VIX feature a correlation of 80% during the sample period considered, we do not simultaneously include them in the multivariate regressions in Table 1.
We use proxies for financial intermediaries constraints proposed in He, Kelly, and Manela (2017): (i) the intermediary capital ratio, which is the aggregate value of market equity divided by aggregate market equity plus aggregate book debt of primary dealers, (ii) the intermediary capital risk factor, or the shock to the capital ratio converted to a growth rate, and (iii) the intermediary leverage ratio squared, or the squared reciprocal of the capital ratio of the intermediary sector. We find negative, but insignificant correlation between the degree of residual mispricing and the intermediary capital ratio, both domestically and internationally. In contrast, there is positive and statistically significant correlation between RMP and the intermediary leverage ratio squared (Figure 2 bottom left panel). This result is not surprising, as RMP itself represents a (nonlinear) correction to standard (linear) asset pricing models that ensures absence of arbitrage opportunities in every state of the world. Moreover, following the predictions of the dynamic intermediary asset pricing model of He and Krishnamurthy (2013) featuring a nonlinear, or state-dependent, association between the asset risk premium and the degree of financial sector turmoil, this relation provides further support that RMP can be interpreted as a measure of (international) financial distress.

To examine whether the intermediary leverage squared is a potential driver of RMP, we run regressions similar in spirit to (14),

\[ RMP_{i,t} = \alpha_i + \beta_i \times \frac{1}{\xi^2_t} + \epsilon_{it}, \]  

with \( i \) denoting the domestic (12), or international market (13), and \( \xi^2 \) the squared intermediary capital ratio. Results are reported in Table 1. The coefficient estimates are positive and significant, suggesting that high leverage (low intermediary capital ratio) states correspond to higher degrees of residual mispricing. Still, the explanatory power is rather mild, around 5%.

Having established the link with financial intermediaries’ capital constraints, we explore next the relation with funding liquidity constraints, as measured by the TED spread, the difference between three-month interbank (LIBOR) rate and U.S. Treasury bills. The spread widens precisely during times of uncertainty, as banks then generally increase interest rates on unsecured loans. Figure 2 (bottom right panel) plots the time-series of U.S. RMP and the TED spread: the positive relation is apparent, suggesting that the price for insuring against bad states is naturally higher during high uncertainty periods, when the TED spread becomes larger. Regression results reported in Table 1 strengthen this evidence. In summary, we show that RMP exhibits

---

21Importantly, the linear component of the extracted SDFs is strongly negatively correlated with the intermediary capital risk factor, and significantly positively correlated with the leverage factor of Adrian, Etula, and Muir (2014).
a close link with market-wide funding liquidity shocks, as measured by the TED spread and capturing a commonality in the financial fragility affecting speculators’ capital and margin requirements. Overall, residual mispricing is high during periods of uncertainty and during episodes of low (funding) liquidity. Since existing proxies explain less than 30% of the time series variation in RMP, we conclude that the latter is a non-redundant measure of financial distress, containing additional relevant information.

A natural question arises whether already the linear part of the SDF ($M^*$) is driven by financial uncertainty and funding constraints proxies, rather than RMP alone, since $M^*$ and its price are likely to increase during periods of financial distress, when the marginal utility of investors is high. To address this possibility, we provide empirical results from regressing $\pi$, the profit from the linear part of the SDF defined in (9), on various determinants in Table 2. The coefficient estimates and the $R^2$ are significantly lower than the ones in Table 1. In multivariate regressions, most of the explanatory power comes from VIX, but also from the financial uncertainty index and TED spread, albeit weaker. We conclude therefore that nonlinear corrections, ensuring arbitrage-free international markets, represent a non-redundant measure of distress, highlighting the importance of residual mispricing for financial markets and the macro-economy.

Table 2: Determinants and linear component of SDFs

This table reports the estimated coefficients from regressing the profit from the linear component of SDF $M^*$ in domestic and international markets on various determinants: $\pi_{i,t} = \alpha_i + \beta_i X_t + \epsilon_{i,t}$. Proxies for financial uncertainty are the financial uncertainty index of Jurado, Ludvigson, and Ng (2015) and the VIX, intermediary leverage ratio squared is from He, Kelly, and Manela (2017), whereas TED spread is from FRED. All variables are standardized. The domestic market is the U.S., whereas international is the average across international markets excluding the U.S. Data are monthly and span January 1985 to June 2020 (January 1986 to June 2020 for TED spread). Newey-West standard errors are reported in brackets. ***", **", and " indicate significant at the 1%, 5% and 10% level, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Panel A: U.S. linear model</th>
<th>Panel B: International linear model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Uncertainty Index</td>
<td>0.193***</td>
<td>0.129*</td>
</tr>
<tr>
<td></td>
<td>[0.088]</td>
<td>[0.069]</td>
</tr>
<tr>
<td>VIX</td>
<td>0.395***</td>
<td>0.391***</td>
</tr>
<tr>
<td></td>
<td>[0.067]</td>
<td>[0.069]</td>
</tr>
<tr>
<td>Intermediary leverage squared</td>
<td>0.135**</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>[0.065]</td>
<td>[0.055]</td>
</tr>
<tr>
<td>TED Spread</td>
<td>0.209***</td>
<td>0.158**</td>
</tr>
<tr>
<td></td>
<td>[0.077]</td>
<td>[0.066]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>3.5%</td>
<td>15.4%</td>
</tr>
</tbody>
</table>

To examine the robustness of our empirical findings, we additionally perform an out-of-sample exercise. We train our model from January 1975 to December 1984, with monthly sequential increments, such that all prior information is accounted for, but free of look-ahead bias. We consider the out-of-sample period starting from January 1985 and ending in June 2020. To study the link between residual mispricing and potential determinants out-of-sample, we run regressions in the spirit of Equation (14) and provide the results in Table 3. The $R^2$s vary between 5% and 10% in univariate...
regressions, and are as high as 16% in multivariate specifications, with most of the out-of-sample explanatory power coming from the financial uncertainty index. Hence, the RMP indices are confirmed as robust and appropriate proxies to measure financial uncertainty, as the evidence provided is not likely to be spurious or driven by overfitting in sample. Moreover, the advantage of our RMP index is that it can be easily computed and is readily available directly from international asset returns.

Table 3: Determinants of RMP (out-of-sample)

This table reports the estimated coefficients from regressing the out-of-sample RMP index on various determinants, in the \( i = \) domestic (12), and international market (13): \( RMP_{i,t} = \alpha_i + \beta_iX_t + \epsilon_{i,t} \). Proxies for financial uncertainty are the financial uncertainty index of Jurado, Ludvigson, and Ng (2015) and the VIX; intermediary leverage ratio squared is from He, Kelly, and Manela (2017), whereas TED spread is from FRED. All variables are standardized. The domestic market is the U.S., whereas international is the average across international markets excluding the U.S. The training period is from January 1975 to December 1984, with monthly sequentially increments, such that all prior information is accounted for. Out-of-sample period is from January 1985 to June 2020 (January 1986 to June 2020 for TED spread). Newey-West standard errors are reported in brackets. ** and * indicate significant at the 5% and 10% level, respectively.

<table>
<thead>
<tr>
<th>Panel A: U.S. Residual Mispricing</th>
<th>Panel B: International Residual Mispricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Uncertainty Index</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>[0.169]</td>
</tr>
<tr>
<td>VIX</td>
<td>0.318</td>
</tr>
<tr>
<td></td>
<td>[0.255]</td>
</tr>
<tr>
<td>Intermediary leverage squared</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>[0.118]</td>
</tr>
<tr>
<td>TED Spread</td>
<td>0.313</td>
</tr>
<tr>
<td></td>
<td>[0.230]</td>
</tr>
<tr>
<td></td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>[0.099]</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>5.3%</td>
</tr>
<tr>
<td></td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td>11.6%</td>
</tr>
</tbody>
</table>

We end this section by investigating the relation between RMP indices in domestic and international markets, and the macro economy, by studying the economic policy uncertainty (EPU) index and the equity market volatility (EMV) index of Baker, Bloom, and Davis (2016) and Baker, Bloom, Davis, and Kost (2019), as well as the Chicago Fed National Activity Index (CFNAI). EPU is constructed to account for (i) newspaper coverage of policy-related economic uncertainty, (ii) the number of federal tax code provisions set to expire in future years and (iii) disagreement among economic forecasters as a proxy for uncertainty. EMV is similar to a newspaper-based Equity Market Volatility tracker that moves with the CBOE Volatility Index (VIX) and with the realized volatility of returns on the S&P 500. EMV is further decomposed into policy-related EMV trackers and a suite of trackers that quantify the importance of each category in the level of U.S. stock market volatility and its movements over time. The CFNAI index is designed to gauge overall economic activity and related inflationary pressure. The results from regressing RMP on different measures of economic uncertainty and activity are reported in Table 4. We document a strong link between RMP and EPU, as the coefficient is positive and significant at the 1% level. The coefficient for CFNAI is not only statistically significant, but
Table 4: RMP, economic policy uncertainty and equity market volatility

This table reports the estimated coefficients from regressing RMP on measures of economic policy uncertainty and equity market volatility, in the $i =$ domestic (12) market (Panel A), and international market (13) (Panel B): $RMP_{i,t} = \alpha_i + \beta_i X_t + \epsilon_{i,t}$. The explanatory variables in $X$ are from Baker, Bloom, and Davis (2016) and Baker, Bloom, Davis, and Kost (2019). EPU is the economic policy uncertainty, CFNAI is the Chicago Fed National Activity Index measuring overall economic activity, and EMV is the equity market volatility. All variables are standardized. The domestic market is the U.S., whereas international is the average across foreign markets excluding the U.S. Data are monthly and span January 1985 to June 2020. Newey-West standard errors are reported in brackets. *** and ** indicate significant at the 1% and 5% level, respectively.

### Panel A: U.S. Residual Mispricing

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPU</td>
<td>0.251***</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>CFNAI</td>
<td>-0.237***</td>
<td>-0.139***</td>
<td></td>
</tr>
<tr>
<td>EMV</td>
<td>0.419***</td>
<td>0.359***</td>
<td></td>
</tr>
<tr>
<td>Macro News and Outlook</td>
<td>0.412***</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>Financial crises</td>
<td>0.275***</td>
<td>0.099</td>
<td></td>
</tr>
<tr>
<td>Financial regulation</td>
<td>0.373***</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td>Monetary policy</td>
<td>0.274***</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>Fiscal policy</td>
<td>0.407***</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>Commodity markets</td>
<td>0.415***</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>6.1%</td>
<td>5.4%</td>
<td>17.3%</td>
</tr>
<tr>
<td></td>
<td>16.8%</td>
<td>7.4%</td>
<td>13.7%</td>
</tr>
<tr>
<td></td>
<td>7.3%</td>
<td>16.4%</td>
<td>17.0%</td>
</tr>
<tr>
<td></td>
<td>19.7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: International Residual Mispricing

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPU</td>
<td>0.226***</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>CFNAI</td>
<td>-0.226***</td>
<td>-0.133**</td>
<td></td>
</tr>
<tr>
<td>EMV</td>
<td>0.416***</td>
<td>0.367***</td>
<td></td>
</tr>
<tr>
<td>Macro News and Outlook</td>
<td>0.406***</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>Financial crises</td>
<td>0.200**</td>
<td>0.083</td>
<td></td>
</tr>
<tr>
<td>Financial regulation</td>
<td>0.346***</td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td>Monetary policy</td>
<td>0.261**</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>Fiscal policy</td>
<td>0.400***</td>
<td>0.141</td>
<td></td>
</tr>
<tr>
<td>Commodity markets</td>
<td>0.425***</td>
<td>0.138</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>5.0%</td>
<td>4.9%</td>
<td>17.1%</td>
</tr>
<tr>
<td></td>
<td>16.3%</td>
<td>3.8%</td>
<td>11.8%</td>
</tr>
<tr>
<td></td>
<td>6.6%</td>
<td>15.8%</td>
<td>17.9%</td>
</tr>
<tr>
<td></td>
<td>18.9%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
also economically meaningful: a lower output in economic activity, as measured by CFNAI, is accompanied by larger RMP. We additionally uncover a tight link between RMP and EMV, as well as its components. Consequently, periods characterized by higher RMP coincide with periods of heightened uncertainty in macro news and outlook. Policy-related EMV trackers of volatility sources for the financial markets, in particular for financial regulation, and monetary and fiscal policies also display a nontrivial explanatory power for RMP. Lastly, we show that RMP is relevant for other asset classes, such as commodity markets. Similar results are obtained in Panel B, for international RMP. Particularly large explanatory power (17%) comes from equity market volatility (EMV), also in the multivariate regression reported in the last column of Table 4. Overall, our empirical findings display a tight link between RMP and proxies of policy and market uncertainty, suggesting that RMP is a relevant, non-redundant index for financial markets’ ability to correctly price traded assets.

4.5 RMP and arbitrage violations in the data

RMP is constructed to ensure the absence of arbitrage. Here, we investigate whether the RMP index is related to observed statistical arbitrage violations in the data, or potential sources of mispricing, such as liquidity, information, sentiment, or noise. To this end, we provide a comparison with the market dislocation index (MDI) of Pasquariello (2014), that captures arbitrage violations in international markets. While we can not exclude the existence of arbitrage opportunities between certain markets empirically, they may be hard to disentangle from market frictions, trading costs, or limits to arbitrage in general, making data collection difficult. Moreover, riskless arbitrage opportunities will likely never be observed in practice since they are short-lived. As a consequence, the data requirements of our index are substantially reduced in that the RMP can be calculated from a bond and a market return only, with no reference to bid-ask spreads and fees. In addition, following Du, Tepper, and Verdelhan (2018b) we consider covered interest rate parity (CIP) violations observed for the developed markets in our sample. We aggregate the daily data from Du, Im, and Schreger (2018a) at monthly frequency, for different tenors, varying from three months to ten years.

We investigate the explanatory and predictive power of domestic and international RMP for financial market dislocations and CIP violations in Table 5. We find a positive and significant relation between RMP and arbitrage violations in the data, both from MDI and CIP, at different tenors considered. The $R^2$s are rather high in both contemporaneous and predictive specifications. For CIP deviations, the coefficients estimates, as well as the $R^2$s feature a non-monotonic relation, first increasing and
This table reports the estimated coefficients from regressing observed arbitrage violations and financial market dislocations on RMP, in the \( i = \text{domestic} \) (12) market, and international market (13). Panel A reports contemporaneous regressions: \( Y_t = \alpha + \beta RMP_{i,t} + \epsilon_t \), while Panel B reports predictive regressions: \( Y_t = \alpha + \beta RMP_{i,t-1} + \epsilon_t \). The independent variables are the market dislocation index (MDI) of Pasquariello (2014) and covered interest rate parity (CIP) deviations in the data, computed as the average across developed markets, at different horizons. All variables are standardized. The domestic market is the U.S., whereas international is the average across foreign markets excluding the U.S. Data are monthly and span January 1985 to December 2009 for MDI (January 2000 to June 2020 for CIP). Newey-West standard errors are reported in brackets. ***, **, and * indicate significant at the 1%, 5% and 10% level, respectively.

### Panel A: Contemporaneous regressions

<table>
<thead>
<tr>
<th></th>
<th>U.S. Residual Mispricing</th>
<th>International Residual Mispricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDI</td>
<td>CIP 3M</td>
<td>CIP 1Y</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.331*</td>
<td>0.320*</td>
</tr>
<tr>
<td></td>
<td>[0.171]</td>
<td>[0.195]</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>10.7%</td>
<td>9.8%</td>
</tr>
</tbody>
</table>

### Panel B: Predictive regressions

<table>
<thead>
<tr>
<th></th>
<th>U.S. Residual Mispricing</th>
<th>International Residual Mispricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDI</td>
<td>CIP 3M</td>
<td>CIP 1Y</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.301**</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td>[0.139]</td>
<td>[0.173]</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>8.7%</td>
<td>5.0%</td>
</tr>
</tbody>
</table>
then decreasing as a function of the tenor, with the highest attained for the mid (two year) tenor. Out-of-sample results in Table 6 confirm once again the robustness of the construction and estimation of RMP, and its applicability as an economic indicator.

To summarize, we conclude that RMP contains additional information not subsumed by proxies of intermediaries’ constraints or existing deviations from the no-arbitrage condition, such as the covered interest rate parity violation. Constructed as an ex-ante measure consistent with no-arbitrage, RMP explains and predicts financial market dislocations and CIP deviations.

Table 6: RMP and statistical arbitrage in the data (out-of-sample)

This table reports the estimated coefficients from regressing observed arbitrage violations and financial market dislocations on RMP, in the $i =$ domestic (12) market, and international market (13). Panel A reports contemporaneous regressions: $Y_t = \alpha + \beta RMP_{i,t} + \epsilon_t$, while Panel B reports predictive regressions: $Y_t = \alpha + \beta RMP_{i,t-1} + \epsilon_t$. The independent variables are the market dislocation index (MDI) of Pasquariello (2014) and covered interest rate parity (CIP) deviations in the data, computed as the average across developed markets, at different horizons. All variables are standardized. The domestic market is the U.S., whereas international is the average across foreign markets excluding the U.S. The training period is from January 1975 to December 1984, with monthly sequentially increments, such that only prior information is accounted for. Out-of-sample periods span January 1985 to December 2009 for MDI (January 2000 to June 2020 for CIP). Newey-West standard errors are reported in brackets. ***,** and * indicate significant at the 1%, 5% and 10% level, respectively.

### Panel A: Contemporaneous regressions

<table>
<thead>
<tr>
<th></th>
<th>MDI</th>
<th>CIP 3M</th>
<th>CIP 1Y</th>
<th>CIP 2Y</th>
<th>CIP 3Y</th>
<th>CIP 5Y</th>
<th>CIP 7Y</th>
<th>CIP 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.158***</td>
<td>0.345</td>
<td>0.191</td>
<td>0.297**</td>
<td>0.301**</td>
<td>0.256***</td>
<td>0.178**</td>
<td>0.140*</td>
</tr>
<tr>
<td></td>
<td>[0.047]</td>
<td>[0.221]</td>
<td>[0.121]</td>
<td>[0.136]</td>
<td>[0.131]</td>
<td>[0.097]</td>
<td>[0.084]</td>
<td>[0.075]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>2.2%</td>
<td>11.6%</td>
<td>3.3%</td>
<td>8.5%</td>
<td>8.7%</td>
<td>6.2%</td>
<td>2.8%</td>
<td>1.6%</td>
</tr>
</tbody>
</table>

### International Residual Mispricing

<table>
<thead>
<tr>
<th></th>
<th>MDI</th>
<th>CIP 3M</th>
<th>CIP 1Y</th>
<th>CIP 2Y</th>
<th>CIP 3Y</th>
<th>CIP 5Y</th>
<th>CIP 7Y</th>
<th>CIP 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.152***</td>
<td>0.296</td>
<td>0.192</td>
<td>0.275*</td>
<td>0.272*</td>
<td>0.228*</td>
<td>0.159</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>[0.058]</td>
<td>[0.217]</td>
<td>[0.138]</td>
<td>[0.150]</td>
<td>[0.144]</td>
<td>[0.120]</td>
<td>[0.109]</td>
<td>[0.084]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>2.0%</td>
<td>8.4%</td>
<td>3.3%</td>
<td>7.2%</td>
<td>7.0%</td>
<td>4.8%</td>
<td>2.1%</td>
<td>0.8%</td>
</tr>
</tbody>
</table>

### Panel B: Predictive regressions

<table>
<thead>
<tr>
<th></th>
<th>MDI</th>
<th>CIP 3M</th>
<th>CIP 1Y</th>
<th>CIP 2Y</th>
<th>CIP 3Y</th>
<th>CIP 5Y</th>
<th>CIP 7Y</th>
<th>CIP 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.137***</td>
<td>0.285</td>
<td>0.268*</td>
<td>0.356**</td>
<td>0.355**</td>
<td>0.303***</td>
<td>0.222**</td>
<td>0.163**</td>
</tr>
<tr>
<td></td>
<td>[0.044]</td>
<td>[0.186]</td>
<td>[0.143]</td>
<td>[0.155]</td>
<td>[0.149]</td>
<td>[0.109]</td>
<td>[0.096]</td>
<td>[0.074]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>1.6%</td>
<td>7.8%</td>
<td>6.8%</td>
<td>12.3%</td>
<td>12.3%</td>
<td>8.8%</td>
<td>4.5%</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

### International Residual Mispricing

<table>
<thead>
<tr>
<th></th>
<th>MDI</th>
<th>CIP 3M</th>
<th>CIP 1Y</th>
<th>CIP 2Y</th>
<th>CIP 3Y</th>
<th>CIP 5Y</th>
<th>CIP 7Y</th>
<th>CIP 10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.135**</td>
<td>0.257</td>
<td>0.266*</td>
<td>0.337**</td>
<td>0.323**</td>
<td>0.273**</td>
<td>0.171</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>[0.053]</td>
<td>[0.180]</td>
<td>[0.158]</td>
<td>[0.167]</td>
<td>[0.162]</td>
<td>[0.132]</td>
<td>[0.112]</td>
<td>[0.083]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>1.5%</td>
<td>6.2%</td>
<td>6.7%</td>
<td>11.0%</td>
<td>10.1%</td>
<td>7.1%</td>
<td>2.5%</td>
<td>1.1%</td>
</tr>
</tbody>
</table>
Having investigated the market forces that might be related to RMP, we next investigate its relation to risk premia. Since the RMP index is by construction orthogonal to asset returns from which it is constructed and therefore unspanned by them, we are interested in uncovering whether this unspanned risk is instead related to asset risk premia more broadly. A long-standing literature on unspanned risk premia documents the possibility that expected returns accommodate information that is not contained in prices (Collin-Dufresne and Goldstein, 2002; Duffee, 2011). As a measure of mispricing, the RMP index lends itself to an investigation in this context. We consider the empirical specification

$$\mathbb{E}[r_{j,t+1} \mid \mathcal{F}_t] = \alpha_j + \beta_j RMP_{dom,t} + \varepsilon_{j,t},$$

(16)

where $\mathbb{E}[r_{j,t+1} \mid \mathcal{F}_t]$ is the conditional expected return on the test assets $j$ at time $t$ for this purpose. The test assets considered are: (i) U.S. equity and international equity converted to U.S. dollars, and (ii) currency returns, i.e. borrowing at foreign risk-free rates and investing in the domestic U.S. risk-free rate. The novelty of our approach stems from the fact that we can use conditional expected returns as the dependent variables, as opposed to realized returns previously employed in the literature.

We estimate the model in Equation (16) using the two-stage procedure of Fama and MacBeth (1973). The methodology proceeds in two steps. First, we run time-series regressions to obtain the betas, or loadings, on RMP. Second, we run cross-sectional regressions on the estimated betas from the first step in order to retrieve the price of risk of RMP, or the $\lambda$s from

$$\mathbb{E}[r_{j,t+1} \mid \mathcal{F}_t] = \lambda_0 + \lambda_j \hat{\beta}_j + \eta_j.$$

(17)

When performing this two-stage regression, we adjust the standard errors to account for errors-in-variables following Shanken (1992), since the betas are estimated in the first step, for heteroskedasticity, as the variance of residuals is not constant, and for potential autocorrelation in error terms.\footnote{Specifically, the standard errors are going to be generalized method of moments (GMM) errors, that are adjusted for serial correlation using Newey and West (1987).}

We provide estimates of the model parameters in (17) for different asset classes in Table 7. We find that RMP explains the cross-section of (international) equities, as well as the cross-section of currencies, with estimated prices of residual mispricing risk that are positive and statistically significant at the 1% level, and $R^2$ of 41% and 90%, respectively. The positive sign resembles the risk compensation for realized volatility...
and realized kurtosis (notably both even moments) in the cross section of US equities (Amaya, Christoffersen, Jacobs, and Vasquez (2015)).

Table 7: RMP index and asset returns

This table reports the prices of risk $\lambda$, Fama and MacBeth t-stats in parentheses, root mean squared errors (RMSE), and the cross-sectional $R^2$ for the U.S. residual mispricing index. Each column contains the following test assets: (i) U.S. equity and international equity converted in U.S. dollars, and (ii) the currency returns, i.e. borrowing at foreign risk-free rates and investing in the domestic U.S. risk-free rate. The first column has eight test assets (individual economies), whereas currencies have seven test assets (corresponding to the bilateral pairs having U.S. as domestic economy). The sample period is January 1985 to June 2020. The standard errors in the Fama and MacBeth approach account for errors-in-variables, heteroskedasticity and autocorrelation in error terms. *** indicates significant at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>Equity</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>$0.035^{***}$</td>
<td>$0.040^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(5.233)$</td>
<td>$(5.077)$</td>
</tr>
<tr>
<td>RMSE</td>
<td>$0.0008$</td>
<td>$0.0003$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$40.8%$</td>
<td>$90.1%$</td>
</tr>
</tbody>
</table>

Figure 3, Panel A, substantiates the regression results by plotting the conditional expected return against the model-implied predicted return. The test assets lie closely to the 45 degree line, suggesting that residual mispricing explains the conditional expected returns considered, especially when considering currencies as test assets. This result is not surprising, as previous evidence documented that crash risk is prevalent in currency markets (see, e.g. Brunnermeier, Nagel, and Pedersen (2008), among others), and thus it is natural that nonlinearities are more important in such cases. In particular, since higher yield currencies exhibit more volatile RMPs, they earn a higher expected return on average (Panel A, right plot). To substantiate the robustness of RMP as a priced risk factor, Table 9 shows the corresponding out-of-sample results. They confirm both the significance and the sign of the in-sample price of risks. We find that in-sample and out-of-sample U.S. RMPs feature a strong positive correlation, of 62%. The greater magnitude of the prices of risk likely reflects the noise stemming from the out-of-sample estimation.

Overall, it follows that investors are aware of the residual structure left by linear models, such that RMP is a relevant risk factor, and assets that are more exposed to this risk seem to require higher conditional expected returns. Consequently, the RMP fits the cross-section of international returns well, with small pricing errors. Importantly, the analysis in this section serves as a validation exercise, and does not try to establish that RMP is the only relevant risk factor driving the cross section of assets. In the robustness Section 5.3, we additionally account for linear factors and we
Figure 3: RMP and expected returns

Panel A: RMP and conditional expectations of returns

Panel B: RMP and sample averages of realized returns

This figure plots in Panel A mean conditional expected returns versus the model-implied predicted mean returns (in percentage) for the domestic RMP using the Fama MacBeth approach. The test assets depicted are the U.S. and international aggregate equity returns, and the currency returns. Panel B plots instead sample averages of realized returns versus the model-implied predicted mean returns (in percentage), for the same test assets as in Panel A. The sample period is January 1985 to June 2020.
find that RMP remains significant also in such cases, in particular in the cross-section of currencies.

To emphasize the relevance of conditional expectations, we perform the same exercise on realized returns, using simple averages. We show in Table 8 that the price of risk is negative, albeit insignificant, when using unconditional expectations of returns. This result is in line with previous evidence documenting the relation with idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang (2006)). Moreover, as illustrated in Figure 3, Panel B, the cross-sectional fit generally worsens.

Table 8: RMP and average (realized) returns

<table>
<thead>
<tr>
<th>Equity</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>((-0.63))</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0013</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>16.2%</td>
</tr>
</tbody>
</table>

Summarizing, our findings suggest that RMP is a priced risk factor, relevant for understanding the cross-section of assets’ conditional mean expected returns, especially those exposed to higher nonlinearities, or crash risk, such as currencies. Moreover, RMP is theoretically and economically motivated as an index insuring absence of arbitrage opportunities.

4.7 RMP and Exchange Rates

With RMP indices for multiple currencies, it is natural to investigate their exchange rate implications. Canonical models in international finance usually assume that markets are complete and integrated (see, e.g., Colacito and Croce (2011), Colacito and Croce (2013), among others). Whenever these assumptions hold, the change in the exchange rate is equal to the ratio of foreign to domestic SDFs, \( X = M_f/M_d \), i.e. the so-called asset market view holds. In more general settings, in which the completeness assumption is relaxed, deviations from the asset market view can be expressed through
Table 9: RMP index and asset returns (out-of-sample)

This table reports the prices of risk $\lambda$, Fama and MacBeth t-stats in parentheses, root mean squared errors (RMSE), and the cross-sectional $R^2$ for the U.S. residual mispricing index. Each column contains the following test assets: (i) U.S. equity and international equity converted in U.S. dollars, and (ii) the currency returns, i.e. borrowing at foreign risk-free rates and investing in the domestic U.S. risk-free rate. The first column has eight test assets (individual economies), whereas currencies have seven test assets (corresponding to the bilateral pairs having U.S. as domestic economy). The training period is from January 1975 to December 1984, with monthly sequentially increments, such that all prior information is accounted for. Out-of-sample period is from January 1985 to June 2020. The standard errors in the Fama and MacBeth approach account for errors-in-variables, heteroskedasticity and autocorrelation in error terms. *** indicates significance at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>Equity</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>0.323***</td>
<td>0.261***</td>
</tr>
<tr>
<td></td>
<td>(5.209)</td>
<td>(5.153)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0008</td>
<td>0.0003</td>
</tr>
<tr>
<td>$R^2$</td>
<td>40.8%</td>
<td>91.6%</td>
</tr>
</tbody>
</table>

a stochastic exchange rate wedge (see, e.g., Backus, Foresi, and Telmer, 2001),

$$X = \frac{M_f}{M_d} \exp(\eta),$$  \hspace{1cm} (18)

where the stochastic wedge $\eta$ measures the degree of international market incompleteness. It is important to mention that in incomplete markets, each particular choice of SDFs will entail a different wedge.

The RMP indices that we derive are consistent with arbitrage-free local markets, in the sense that each economy has its own SDF, and measure of the correction to linear asset pricing models. We study next their role in international settings. To this end, we rewrite Equation (18) as

$$X = \frac{M_f^*}{M_d^*} \frac{(1 + M_f^*/M_f^* \exp(\eta))}{(1 + M_d^*/M_d^*)} \text{exp(\eta)} = \frac{M_f^*}{M_d^*} \exp(\tilde{\eta}),$$  \hspace{1cm} (19)

with $\tilde{\eta}$ the stochastic wedge implied by the ratio of the linear SDF projections $M_f^*$ and $M_d^*$. Sandulescu, Trojani, and Vedolin (2021) provide an economic interpretation of the stochastic wedge as a measure of unspanned risk, capturing potential nonlinearities in each country. We further decompose this wedge to reflect the no-arbitrage correction, or the projection error on the space of linear returns with the minimal variance, i.e. the

---

23 We consider in this case a multiplicative wedge, similar to Lustig and Verdelhan (2019). An additive wedge, i.e. $X = \frac{M_f}{M_d} + \lambda$, following Bakshi, Cerrato, and Crosby (2018), delivers qualitatively similar implications. We obtain a correlation of 0.98 between the multiplicative and additive wedges.
term in the middle in the above equation, capturing the relative residual mispricing. Notice that whenever markets are complete, $\eta$ is naturally equal to zero. Interestingly, if the relative importance of $M^\circ$ is similar in the two markets, such that the ratio is close to one, we retrieve the standard stochastic wedge relative to the linear minimum variance SDFs. Moreover, by construction, since $M^\circ$ is orthogonal to $M^*$ in each market, it will have no effect on the prices of assets. However, it is an empirical question whether the ratio will impact exchange rates in international markets settings.

**Figure 4: RMP and Exchange Rates**

![Panel A: Relative residual mispricing in international markets](image)

![Panel B: Stochastic wedge in international markets](image)

This figure plots the time series of the relative RMP term $\frac{(1+M^\circ_f/M^\circ_d)}{(1+M^*_{f,d})}$ (Panel A) and the time series of the stochastic wedge $\tilde{\eta}$, relative to the linear HJ SDFs (Panel B). Gray bars denote NBER recessions. The sample period is January 1985 to June 2020.

Figure 4 shows that the ratio of relative RMP is generally close to one. It follows that the correction, or relative mispricing in each market is roughly proportional, such that their ratio in (19) does not contribute to the change in exchange rates, similarly to the wedge $\eta$. This pattern is further depicted in Figure 5, where we report the $R^2$ from regressing the changes in exchange rates on the ratio $\frac{M^\circ_f}{M^\circ_d}$, as well as the $R^2$ from
regressing the changes in exchange rates on the ratio of RMPs. Consequently, the latter provides no explanatory power for exchange rates. Hence, when international financial markets are integrated, the risk factors driving premia are similar, such that the residual mispricing is proportional in domestic and foreign markets. Our findings are consistent with Jensen, Kelly, and Pedersen (2021), who document that OLS alphas for U.S. factors are very similar to their international counterpart.

Figure 5: Exchange rate fit

This figure plots the $R^2$ from regressing the changes in exchange rates on the ratio of linear HJ SDFs, as well the $R^2$ from regressing the changes in exchange rates on the ratio of RMPs. The sample period is January 1985 to June 2020.

Still, Figure 4 shows relevant spikes around periods of financial distress and recessions, such as during the 1987 crash, the Asian crisis and LTCM default, the financial crisis of 2008 and more importantly, during the recent COVID-19 crisis. Overall, most of the spikes encountered are below one, suggesting that during financial turmoil, on average, the degree of residual mispricing is more relevant in the domestic U.S. market, rather than in the foreign economies studied. Overall, we find a positive cross-market correlation between the relative RMPs across the economies studied.24

Lastly, Panel B of Figure 4 plots the time series of the stochastic wedge $\tilde{\eta}$ defined in Equation (19). Accordingly, the international unspanned risk fluctuates around zero, exhibiting both positive and negative spikes, especially during financial distress periods.

Next, we explore whether the stochastic wedge is related to asset risk premia. We have established in Section 4.6 that RMP is a priced risk factor in the cross section of (international) equities and currencies. We report in Table 10 the results obtained following the same Fama MacBeth approach as in Section 4.6, where we consider the stochastic wedge ($\eta$), and the stochastic wedge ($\tilde{\eta}$) in Equation 19.

24We report the cross-market correlations between the relative RMPs in Appendix Table C2.
respectively, as factors. We plot the estimated prices of risk in Table 10. There are two noteworthy observations. First, even if the stochastic wedge appears to be priced in the cross-section of international equities and currencies, the reported $R^2$ are significantly lower than the ones in Table 7 having RMP as a risk factor. Second, it is crucial what SDFs are used to determine the stochastic wedge, as they will give rise to different implications in terms of the associated prices of risk. Using the SDFs that ensure absence of arbitrage opportunities (Panel A), the corresponding wedge naturally commands a positive risk premia for international equities and currencies. By using instead the ratio of the minimum variance SDFs implied by linear models, that do not rule out arbitrage (Panel B), the ensuing wedge gives rise to distinct implications: it commands a negative price of risk for currencies, but a positive one for international equities. Moreover, the $R^2$s are lower than the ones in Panel A of Table 10. Overall, we conclude that using SDFs accounting for residual mispricing yields more robust and stable results than using linear asset pricing models alone.

**Table 10: Stochastic wedge and asset returns**

This table reports the prices of risk, Fama and MacBeth t-stats in parentheses, root mean squared errors (RMSE), and the cross-sectional $R^2$ for the stochastic wedge. Panel A (B) reports results for $\eta$ ($\tilde{\eta}$). The test assets considered are: (i) the U.S. and international equity converted in U.S. dollars, and (ii) the currency returns, i.e. borrowing at foreign risk-free rates and investing in the domestic U.S. risk-free rate. The first column has eight test assets (individual economies), whereas the currency has seven test assets (corresponding to the bilateral pairs having U.S. as domestic economy). The sample period is January 1985 to June 2020. The standard errors in the Fama and MacBeth approach account for errors-in-variables, heteroskedasticity and autocorrelation in error terms. *** indicates significant at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>Equity</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Stochastic wedge ($\eta$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.214***</td>
<td>0.229***</td>
</tr>
<tr>
<td></td>
<td>(4.879)</td>
<td>(3.635)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0008</td>
<td>0.0009</td>
</tr>
<tr>
<td>$R^2$</td>
<td>32.22%</td>
<td>11.22%</td>
</tr>
<tr>
<td>Panel B: Stochastic wedge implied by linear models ($\tilde{\eta}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.281***</td>
<td>-0.167***</td>
</tr>
<tr>
<td></td>
<td>(5.271)</td>
<td>(-3.083)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0008</td>
<td>0.0010</td>
</tr>
<tr>
<td>$R^2$</td>
<td>29.02%</td>
<td>4.79%</td>
</tr>
</tbody>
</table>
5 Robustness

In this section, we perform several additional analyses that further our understanding of RMP. In Section 5.1 we investigate how option-type returns alter the properties of RMP, Section 5.2 confronts RMP with the mispricing factors from Stambaugh and Yuan (2017), and Section 5.3 corroborates the empirical success of RMP as a driver for expected returns. Section 5.4 processes RMP through an international benchmark model.

5.1 Hedging in nonlinear instruments and RMP

In this section we investigate how RMP behaves if nonlinear trading strategies become available to investors. Specifically, we are interested in the question how the behaviour of \( M_{t+1}^o \) is altered when adding nonlinear securities to the linear SDF. In the context of our setting, we approach this problem by adding \( r_{t+1}^2 \), the excess return squared, as an additional term to \( M_{t+1}^* \). This instrument corresponds to the payoff of a Martin (2017) simple variance swap. The price of this payoff is an option portfolio holding out-of-the-money calls and out-of-the-money puts scaled by the squared forward price. A lower bound for the price of this instrument is the price of an at-the-money straddle scaled by the squared forward price of the underlying. Denote this lower bound at time \( t \) by \( C_t \). To incorporate this pricing information, we add to the original problem (5) the constraint that

\[
E^P \left[ M_{t+1} r_{t+1}^2 \mid \mathcal{F}_t \right] \geq C_t. \tag{20} 
\]

This lower bound induces \( M_{t+1}^* \) to have a linear term in \( r_{t+1}^2 \), meaning that the linear benchmark model \( M_{t+1}^* \) accommodates an additional quadratic hedging term. Note that now already \( M^* \), as a quadratic polynomial, could become a valid SDF.

Empirically, we need to work with observed option prices to implement constraint (20). Since options data for our international sample outlined in Section 4 are not readily available, we restrict ourselves to the S&P 500 index return, along with S&P 500 options data from OptionMetrics for this exercise. Hence, we derive RMP from Equation (5) only for the U.S. market with the additional constraint (20), and plot in Figure 6 the corresponding time series, along with the original RMP extracted from the linear model. The two series are termed RMP, and variance hedged RMP, accordingly.

The scale of variance hedged RMP is orders of magnitude smaller than the one of the original RMP, suggesting that nonlinear hedging in the options market will indeed reduce residual mispricing. As another striking observation, the variance hedged RMP does not show any extraordinary movement during recessions. This fact may suggest that trading variance contracts is particularly valuable during distress periods. Finally,
This figure plots the time series of the U.S. RMP, as well as the variance hedged U.S. RMP (dashed line). Gray bars denote NBER recessions. The data are S&P 500 returns, as well as at-the-forward put and call options. The sample period is January 1996 to December 2019.

the few times the variance hedged RMP does spike, it goes in the opposite direction of the original RMP.

Overall, we find that augmenting the set of assets available to investors to account for nonlinear payoffs decreases the magnitude of the underlying residual mispricing.

5.2 Mispricing factors

In this section, we explore the link between our residual mispricing index and existing factors capturing mispricing in the U.S. equity market. Specifically, Stambaugh and Yuan (2017) construct two mispricing factors by aggregating the information contained in 11 prominent anomalies, i.e. long-short decile portfolios of U.S. equities, sorted on different firm characteristics, related to management and performance. We report in Table 11 the correlation between our RMP indices and the mispricing factors. We find no significant relation, suggesting that we uncover a relevant nonlinear factor, whereas Panel B shows that the mispricing factors are positively correlated with our linear component of the SDFs, both in domestic and international markets.
Table 11: Correlation between RMP and mispricing factors

This table reports the correlation between RMP (linear component of SDF) and the mispricing factors of Stambaugh and Yuan (2017) in Panel A (Panel B). MGMT and PERF are mispricing factors aggregating information from 11 anomalies, related to firm management and performance. The domestic market is the U.S., whereas international is the average across international markets excluding the U.S. Data are monthly and span January 1985 to December 2016. *** indicates significant at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>Panel A: Residual mispricing</th>
<th>Panel B: Linear component of SDF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U.S.</td>
<td>International</td>
</tr>
<tr>
<td>MGMT</td>
<td>0.051</td>
<td>0.062</td>
</tr>
<tr>
<td>PERF</td>
<td>0.005</td>
<td>0.006</td>
</tr>
</tbody>
</table>

5.3 RMP, linear component of SDF and expected returns

Since international equity and currency premia may be related to other priced risk factors than RMP, we investigate this possibility by including in specification (16) of Section 4.6 the linear component of the SDF. Specifically, we consider

$$\mathbb{E}[r_{j,t+1} | \mathcal{F}_t] = \alpha_j + \beta_j RMP_{dom,t} + \gamma_j \pi_{dom,t} + \varepsilon_{j,t}, \quad (21)$$

and

$$\mathbb{E}[r_{j,t+1} | \mathcal{F}_t] = \lambda_0 + \lambda_1 \hat{\beta}_j + \lambda_2 \hat{\gamma}_j + \eta_j. \quad (22)$$

We report the estimated prices of risk in Table 12. In general, the price of risk for RMP remains positive and statistically significant, even after accounting for the linear factors embedded in $\pi_{dom}$. Notice that, by construction, $M^*$ will price the corresponding bilateral assets from which it is derived. Interestingly, when considering average domestic RMP in (12), and examining its performance across different test assets than the ones from which it is constructed, we find that $\pi$ is relevant for the cross-section of international equities and currencies. Domestic RMP remains relevant, in particular for currencies.

Overall, our findings suggest that RMP is a relevant non-redundant risk factor, as it is not entirely subsumed by linear factors.
Table 12: RMP, π, and asset returns

This table reports the prices of risk $\lambda_1$ and $\lambda_2$ for the domestic RMP and $\pi$, Fama and MacBeth $t$-stats in parentheses, root mean squared errors (RMSE), and the adjusted cross-sectional $R^2$. The test assets considered are: (i) U.S. equity and international equity converted in U.S. dollars, and (ii) the currency returns, i.e. borrowing at foreign risk-free rates and investing in the domestic U.S. risk-free rate. The first column has eight test assets (individual economies), whereas currencies have seven test assets (corresponding to the bilateral pairs having U.S. as domestic economy). The sample period is January 1985 to June 2020. The standard errors in the Fama and MacBeth approach account for errors-in-variables, heteroskedasticity and autocorrelation in error terms. ***, and ** indicate significant at the 1%, and 5% level, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Equity</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>0.020</td>
<td>0.050***</td>
</tr>
<tr>
<td></td>
<td>(1.436)</td>
<td>(5.504)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.329</td>
<td>-0.293**</td>
</tr>
<tr>
<td></td>
<td>(1.358)</td>
<td>(-2.253)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0008</td>
<td>0.0003</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>33.2%</td>
<td>89.6%</td>
</tr>
</tbody>
</table>

Figure 7: RMP, π, and expected returns

This figure plots mean conditional expected returns versus the model-implied predicted mean returns (in percentage) for the domestic RMP and the linear component of the domestic SDF using the Fama MacBeth approach. The test assets depicted are the U.S. and international aggregate equity, and currency returns. The sample period is January 1985 to June 2020.
5.4 RMP in Benchmark Models

RMP is computed nonparametrically from moments of input distributions having finite second-order moments. It can be interpreted as an arbitrage insurance, as it is the minimal-cost nonlinear trading strategy of a representative agent optimizing her Sharpe ratio portfolio, such that markets are arbitrage free. Since the input is a moment matrix (A.3), it can also originate from parametric models. In the following, we illustrate the behavior of RMP using the well-known benchmark model from Colacito and Croce (2013) using the Gaussian approximation developed in their Appendix C.

The ingredients are dynamics for the log SDF

\[ m^i_{t+1} = \log \delta - \frac{1}{\psi} \Delta c^i_{t+1} + \left( \frac{1}{\psi} - \gamma \right) \log \tilde{U}^i_{t+1} \]

\[ - \frac{1}{\psi - \gamma} \log \mathbb{E} \left[ \exp((1 - \gamma) \log \tilde{U}^i_{t+1})|\mathcal{F}_t \right], \quad i \in \{d, f\}, \]

where \( \Delta c^i_{t+1}, \log \tilde{U}^i_{t+1} \), denoting the log consumption growth and the continuation value of utility, respectively, are assumed to be jointly normal with mean \( \mu = (0, 0, 0, 0)^\top \), and covariance matrix \( \Sigma \). The parameters \( \psi, \gamma \) control risk aversion and preference for early resolution of risk. For utmost parsimony, but without applying log-linearization, we specify the equity returns as \( \log R^i_{1,t+1} = -m^i_{t+1}, \quad i \in \{d, f\} \), so that the Euler conditions are satisfied. Furthermore, following Colacito and Croce (2013), we assume that the log exchange rate growth is given by the log SDF difference \( \log X_{t+1} = m^f_{t+1} - m^d_{t+1} \). The corresponding bond prices are accordingly \( \mathbb{E} \left[ \exp m^i_{t+1}|\mathcal{F}_t \right] \). Consistently with our empirical section, we define as state variables the domestic excess return \( R^d_{0,t+1} - R^d_{0,t+1} \), the excess return associated with the foreign equity return in domestic currency \( R^f_{1,t+1} X_{t+1} - R^d_{0,t+1} \), and the carry \( R^f_{0,t+1} X_{t+1} - R^d_{0,t+1} \). We compute the moments of these state variables using the moment-generating function of consumption growth and long-run factors and follow the procedure outlined in Appendix A to derive the coefficients of \( M^*_{t,t+1} \) and \( M^\circ_{t,t+1} \).

We restrict our attention to the U.S. - UK pair, with U.S. being the domestic, and UK the foreign market, similarly to the calibration used in Colacito and Croce (2013). Figure 8 plots the RMP extracted from the benchmark long run risk model and our nonparametric RMP obtained with unconditional estimates, for different levels of the realized carry trade. There are several noteworthy observations. First, RMP in both settings features a U-shape pattern and is high and positive during distress periods, i.e. whenever both domestic and foreign returns are negative. Second, our nonparametric

\cite{Croce2021} shows that \( \log \tilde{U}^i_{t+1} \) are not exactly long-run shocks, and that the error is reflected in a constant. For simplicity, we ignore this effect here.

\[ 25 \]
This figure shows RMP as a function of carry trade considered (left column in Figure 9), a different picture emerges during calm periods, conditional RMP is similar to the one extracted from benchmark models incorporating long run risk components.

specification for the RMP appears to be more stable when varying the level of the carry trade, whereas RMP in long-run risk models displays a higher variability. Overall, the unconditional nonparametric RMP is similar to the one extracted from benchmark models incorporating long run risk components.

In the following, we highlight the importance of conditioning information and time-varying estimates, by plotting in Figure 9 our conditional RMP. For illustrative purposes, we focus on two polar cases: one in which markets are relatively calm (July 2005) and another one during the collapse of the Lehman Brothers during the financial crisis (September 2008). While during calm periods, conditional RMP features similar properties with the unconditional RMP, regardless of the level of the carry trade considered (left column in Figure 9), a different picture emerges during
This figure shows RMP as a function of \( r^*_i \), where \( r^*_i = \log R^*_i, i \in \{d, f\} \) computed from our conditional estimations for two dates: July 2005 (left column) chosen to represent a calm period, and September 2008 (right column), reflecting the collapse of Lehman Brothers during the financial crisis. The range of \( r^*_i \) and \( r^*_f \) is chosen to be two standard deviations around the mean.

distress periods: conditional RMP features an inverted U-shape, such that it is higher on average, positive, and flatter, suggesting that insuring against mispricing becomes more expensive.
6 Conclusion

We develop a measure of mispricing from basic asset returns. It is constructed as the nonlinear stochastic discount factor component with the smallest variance that renders a linear asset pricing model arbitrage-free in all states of the world. Our measure is easy to compute, it has parsimonious data requirements, and it is fully conditional. In particular, we do not need option prices to identify the nonlinear component of asset prices. We term our measure Residual MisPricing (RMP). We show that conditional linear asset pricing models perform well on average (i.e. display low RMP), and during normal times, while they imply a larger (residual) mispricing (high RMP) during crisis periods.

We exploit the weak data requirements of RMP, and produce time series for the United States, United Kingdom, Switzerland, Australia, Canada, Japan, New Zealand, and the Euro area. Empirically, RMP is related to financial uncertainty, VIX, as well as measures of economic policy uncertainty. We show that RMP drives equity and currency risk premia, even after accounting for linear factors. This provides further evidence for unspanned components that cannot be found in prices of basic assets, such as international stocks and bonds. Lastly, RMP is related to nonlinear proxies of intermediary leverage constraints and exhibits a close link with market-wide funding liquidity shocks, as measured by the TED spread. It therefore constitutes a financial distress measure that is not only theoretically motivated from first no-arbitrage principles, but also non-redundant. Accordingly, we find that RMP predicts future market dislocations, including covered interest rate parity (CIP) deviations. Overall, we document that RMP is high during periods of uncertainty and during episodes of low (funding) liquidity, exhibiting a substantial effect not only across different asset classes, but also a close link with macro-economic indicators. We show that augmenting the set of assets available to investors to account for nonlinear payoffs decreases the magnitude of the underlying residual mispricing. Our findings thus contribute to the study of the relation between uncertainty and mispricing.

Our framework is general and can be easily extended, for example with modifications to introduce trading frictions such as short-selling constraints. We leave this, an extensive study involving larger cross sections, and related topics for future research.
References


Koijen, R. S. and M. Yogo (2020): “Exchange rates and asset prices in a global

Maryland.


of Imperial College Press Optimization Series, Imperial College Press.

currency markets and other asset classes,” Journal of Financial Economics, 114,
197–225.

Ludvigson, S. C., S. Ma, and S. Ng (2015): “Uncertainty and business cycles:
exogenous impulse or endogenous response?” Working paper.


Economics, 111, 527–553.

financial markets help to explain exchange rates?” American Economic Review, 109,
2208–44.

Martin, I. (2017): “What is the expected return on the market?” Quarterly Journal


Press.


A Optimization program

From (6), each pricing kernel $M_{i,t+1} = w_{i,t}^\top B_{i,n,t+1}$. A sufficient condition for the constraint $w_{i,t}^\top B_{i,n,t+1} \geq 0$ is the so-called sum-of-squares (SOS) condition

$$w_{i,t}^\top B_{i,n,t+1} = B_{i,n/2,t+1}^\top A_{i,t} B_{i,n/2,t+1}$$

for some symmetric and positive semidefinite matrix $A_{i,t}$. Expanding the right-hand side of (A.1), it can be seen that the coefficients $w_{i,t}$ are related to $A_{i,t}$ through a linear map $L$. Equipped with the constraints that follow from this condition, Program (5) thus becomes

$$\begin{align*}
\text{minimize} & \quad w_{i,t}^\top H_{i,n,t} w_{i,t} \\
\text{subject to} & \quad G_{i,n,t} w_{i,t} = 0_m, \quad w_{i,t}^\top \mu_{i,n,t} = \frac{1}{R_{0,t+1}} \\
& \quad w_{i,t} = L(A_{i,t}), \quad A_{i,t} \succeq 0, \quad \text{with}
\end{align*}$$

$$H_{i,n,t} := \mathbb{E}^\mathbb{P} \left[ B_{i,n,t+1}^\top B_{i,n,t+1} \mid \mathcal{F}_t \right].$$

The vector $\mu_{i,n,t}$ denotes the first row of, and $G_{i,n,t}$ denotes rows 2 through $m_i + 1$ of $H_{i,n,t}$. Since $R_{0,t+1}$ and $H_{i,n,t}$ are $\mathcal{F}_t$-measurable, so are the coefficients $w_{i,t}$. This program is a convex conic optimization problem for which efficient solution algorithms exist that yield an optimal and unique solution (provided the constraints are feasible). For our empirics we use the software package MOSEK. Solutions obtain in fractions of a second. If needed, condition (A.1) could be made strict by a translation with some $\epsilon > 0$ through a change of variables. In our empirical application, the SDF never touches zero, such that in the data the inequality is strict.

B RKHS conditional moments

From the previous section, to solve program (A.2), we need to estimate conditional moments. In particular, we need to estimate matrix $H_{i,n,t}$ that is positive semidefinite (PSD) for population moments, but not necessarily so for estimated moments. To ensure the PSD property of $H_{i,n,t}$ that is also necessary to maintain convexity of program (A.2), we resort to reproducing kernel Hilbert space (RKHS) embedding of conditional distributions with subsequent nearest PSD regularization.

RKHS embedding of conditional distributions is a technique from machine learning to estimate the conditional expectation function (Song et al., 2009). A complete self-contained introduction of this topic is beyond the scope of this paper, so we introduce
very briefly the bare minimum of the necessary concepts. A RKHS $H$ on $\mathcal{X}$ with kernel $k_H$ is a Hilbert space of functions $f : \mathcal{X} \to \mathbb{R}$. Its defining properties are written in terms of its inner product for $x, x' \in \mathcal{X}$,

$$
\langle f, k(x, \cdot) \rangle_H = f(x),
$$

$$
\langle k(x, \cdot), k(x', \cdot) \rangle_H = k(x, x').
$$

The kernel hence is a linear evaluation functional that through the inner product maps function objects to their evaluation at any $x \in \mathcal{X}$. The marvelous power of RKHS arises through the fact that the kernel function $k(x, \cdot)$ itself may be infinite-dimensional, while when used inside the inner product it maps elements of $H$ to their evaluation at a point. This mapping from a potentially infinite-dimensional feature space to simple evaluations is called the kernel trick. Learning about objects $H$ that best represent given observed data, say $x_1, \ldots, x_N$, usually leads to a linear algebra operation involving the symmetric and positive definite kernel matrix

$$
K_H := \begin{pmatrix}
  k(x_1, x_1) & \cdots & k(x_1, x_N) \\
  \vdots & \ddots & \vdots \\
  k(x_N, x_1) & \cdots & k(x_N, x_N)
\end{pmatrix}
$$

through a so-called representer theorem.

### B.1 RKHS embedding of conditional distributions

Turning to the problem of the present paper, RKHS embedding of distributions is performed in the tensor product space $F \otimes G$ of two RKHS $F$ and $G$ (with kernels $k_F$ and $k_G$, respectively), where it has to be assumed that for all $g \in G$ the conditional expectation $\mathbb{E}[g(R) \mid z] \in F$. With this in place we may define $\mu_{R \mid z} := \mathbb{E}[k_G(R, \cdot) \mid z]$ so that

$$
\mathbb{E}[g(R) \mid z] = \langle g, \mu_{R \mid z} \rangle_G.
$$

From Song et al. (2009) and Park and Muandet (2020), given conditioning data $z_1, \ldots, z_N$ and return realizations $R_{s,1}, \ldots, R_{s,N}$, the desired object $\mu_{R \mid z}$ can be estimated with an additional regularization parameter $\lambda$ via

$$
\widehat{\mu_{R \mid z}} = \begin{pmatrix}
  k_G(R_{s,1}, \cdot) & \cdots & k_G(R_{s,N}, \cdot)
\end{pmatrix}
\begin{pmatrix}
  k_F(z_1, \cdot) \\
  \vdots \\
  k_F(z_N, \cdot)
\end{pmatrix}

(K_F + N\lambda I_N)^{-1}
$$

(B.5)
where $I_N$ is the identity matrix of dimension $N$. The regularization parameter $\lambda$ controls overfitting. With this set-up in place, to estimate the first conditional moment of $R$ given some value $z$, it suffices to perform a linear-algebra operation

$$\mathbb{E}\left[R_i \mid z\right] = \left(R_{i,1} \cdots R_{i,N}\right) \left(K_F + N\lambda I_N\right)^{-1} \begin{pmatrix} k_F(z_1, z) \\ \vdots \\ k_F(z_N, z) \end{pmatrix}.$$  

(B.6)

Thus, the conditional expectation is a function of $z$ parameterized by the realized $z_1, \ldots, z_N$.

For our empirical study we use Gaussian kernels

$$k_F(z, z') := e^{-\frac{1}{2} \left\| z - z' \right\|_2^2 \sigma^2}.$$  

(B.7)

The literature on RKHS has found the particular choice of kernel not important for the results. We perform the above operation (B.6) for all the moments making up the matrix $H_{i,n}$ in the optimization program A. Subsequently, to obtain a PSD matrix, we find the closest (in Frobenius norm) positive definite matrix and use this matrix in our optimization program (A.2).

B.2 In-sample and out-of-sample cross validation

We describe here out-of-sample cross validation of the parameters $\sigma > 0$ and $\lambda > 0$, with the former parameterizing the Gaussian kernel (B.7), and the latter controlling the regularization in (B.5). As with traditional local linear and local constant nonparametric regression, also RKHS exhibit a bias-smoothness tradeoff. This is clear from Eq. (B.5) that would yield the observations as conditional expectations with $\lambda = 0$.

To perform cross validation in our context, suppose that we would like to estimate conditional expectations using information up to, and not including some index $N \leq N$. From an initial sample ranging from one to $N := \lceil N/3 \rceil$ we compute the moment predictions

$$\mathbb{E}\left[R_i \mid z_N\right], \ldots, \mathbb{E}\left[R_i \mid z_{\min(N+S,N)}\right]$$

$$= \left(R_{i,1} \cdots R_{i,N}\right) \left(K_F + N\lambda I_N\right)^{-1} \begin{pmatrix} k_F(z_1, z_N) & \cdots & k_F(z_1, z_{\min(N+S,N)}) \\ \vdots & \ddots & \vdots \\ k_F(z_N, z_N) & \cdots & k_F(z_1, z_{\min(N+S,N)}) \end{pmatrix},$$  

(B.8)
for all $\alpha$ powers necessary to fill the matrix $B_{i,n}^T B_{i,n}$ giving $H_{i,n,N}, \ldots, H_{i,n,\min(N+S,N)}$ for our evaluation horizon $S$ ($S$-fold time series cross validation). The sum of the squared prediction errors can be written concisely in Frobenius norm as

$$C_{N,S} := \|B_{i,n,N+1}^T B_{i,n,N+1} - H_{i,n,N}\|^2 + \cdots + \|B_{i,n,\min(N+S+1,N)}^T B_{i,n,\min(N+S+1,N)} - H_{i,n,\min(N+S,N)}\|^2.$$  

Denoting by $\# := \lfloor (N - N)/S \rfloor$ the number of windows used in the cross validation, we

$$\min_{\sigma,\lambda} C_{N,S} + \cdots + C_{N+\#S,S} \quad \text{(B.9)}$$

to find the optimal kernel and regularization parameters $\sigma$ and $\lambda$. Since this problem is non-convex, we obtain a solution from solving a large number of optimization problems started from a random initial parameter vector. The in-sample results are obtained from using this procedure with the entire data set. Our out-of-sample estimates are obtained by performing this procedure for each of the data points at indices $N/3, N/3+1, \ldots, N$, so that the first $N/3$ data points constitute the initial training period.

## C Appendix Figures and Tables

Appendix C contains additional figures and tables.

### Figure C1: Correlations Between Conditional and Unconditional SDFs

This figure shows the scatter plot between conditional and unconditional domestic SDFs, for the corresponding bilateral pair. The sample period is January 1985 to June 2020.
This figure plots the time series of the HJ bound of the domestic SDF corresponding to each bilateral pair studied. Gray bars denote NBER recessions. The sample period is January 1985 to June 2020.

Table C1: Cross-market correlations of the degree of residual mispricing

This table reports the correlations between the residual mispricing indices across different economies. The international RMP is computed as the average across all foreign countries (excluding the U.S.). Data are monthly and span January 1985 to June 2020. *** indicates significant at the 1% level.

<table>
<thead>
<tr>
<th></th>
<th>U.S.</th>
<th>UK</th>
<th>CH</th>
<th>JP</th>
<th>EU</th>
<th>AU</th>
<th>CA</th>
<th>NZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>0.872***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>0.802***</td>
<td>0.857***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU</td>
<td>0.821***</td>
<td>0.778***</td>
<td>0.679***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU</td>
<td>0.782***</td>
<td>0.846***</td>
<td>0.821***</td>
<td>0.663***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>0.837***</td>
<td>0.874***</td>
<td>0.804***</td>
<td>0.727***</td>
<td>0.811***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NZ</td>
<td>0.712***</td>
<td>0.691***</td>
<td>0.607***</td>
<td>0.616***</td>
<td>0.615***</td>
<td>0.682***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>International</td>
<td>0.920***</td>
<td>0.952***</td>
<td>0.900***</td>
<td>0.854***</td>
<td>0.893***</td>
<td>0.918***</td>
<td>0.778***</td>
<td>0.891***</td>
</tr>
</tbody>
</table>

Table C2: Cross-market correlations of the arbitrage-free correction

This table reports the cross-market correlations between relative RMP (19). Data are monthly and span January 1985 to June 2020.

<table>
<thead>
<tr>
<th></th>
<th>CH</th>
<th>JP</th>
<th>EU</th>
<th>AU</th>
<th>CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>0.288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>0.040</td>
<td>0.563</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP</td>
<td>0.340</td>
<td>0.427</td>
<td>0.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EU</td>
<td>0.047</td>
<td>0.165</td>
<td>0.117</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>AU</td>
<td>0.042</td>
<td>0.043</td>
<td>-0.054</td>
<td>0.023</td>
<td>0.044</td>
</tr>
<tr>
<td>CA</td>
<td>0.018</td>
<td>0.054</td>
<td>0.004</td>
<td>0.031</td>
<td>0.042</td>
</tr>
</tbody>
</table>
D Time-varying weights

In this section, we document the time-series properties of the weights, or coefficients, associated with the linear model $M^*$. For illustrative purposes, we focus on one currency pair, namely the U.S. dollar - Swiss franc. Results are qualitatively similar for the remaining of the currency pairs in our sample.

Figure D3: Time-varying Weights

Panel A: Domestic (U.S.) investor’s asset weights

Panel B: Foreign (Swiss) investor’s asset weights

This figure plots the time-varying asset weights (the coefficients of $M^*$) of a domestic (foreign) U.S. (Swiss) investor in Panel A (Panel B). Gray bars denote NBER recessions. The sample period is January 1985 to June 2020.

Figure D3 plots the time-varying weights in the domestic and foreign equities, as well as in the respective currency (left y-axis), and in the risk-free rate (right y-axis). Specifically, these are the optimal weights of a Sharpe ratio portfolio corresponding to the linear minimum variance SDF $M^*$. There are several noteworthy observations. First, weights in risky assets exhibit a significant variation across time, suggesting that analyzing only unconditional models might be an oversimplification. Second,
the currency weights display large swings, that mirror the behavior of the short-term risk-free bond exposure. During crisis periods, there is a flight to safety, as there is a drop in the currency weight followed by a concomitant rise in the domestic bond. Still, the risk-free bond exposure is much less volatile, with an average coefficient of 1.015. Third, the exposure to domestic equity (Panel A) is similar to the one entered in foreign equity by a foreign investor (Panel B). The same relation holds for the foreign (domestic) equity of a domestic (foreign) investor. The exposure to the currency has a natural inverse relation from the perspective of a domestic investor, as opposed to the foreign investor.

To provide further support for the importance of time-varying weights, and subsequently, time-varying risk premia, we plot in Figure D4 conditional versus unconditional, or constant, weights associated with the currency return. Importantly, unconditional weights are virtually the same for domestic and foreign investors, of approximately 1.5 during the sample period analyzed, whereas there are clear differences in the conditional weights. For instance, the sudden adjustment in the weight following the unpeg of the Swiss franc from the Euro at the beginning of 2015 is apparent, but completely ignored by an unconditional approach. Overall, we obtain time-varying coefficients that are plausible and well behaved.

Figure D4: Time-varying currency weights

This figure plots the time-varying currency weights (coefficients of $M^*$) versus the unconditional weight of a domestic (foreign) U.S. (Swiss) investor in the top panel (bottom panel). The red dot during 2015 corresponds to the unpeg of the Swiss franc from the Euro. Gray bars denote NBER recessions. The sample period is January 1985 to June 2020.

Note that from the perspective of a foreign investor, foreign equity is the U.S. equity.