Universität Trier

Quantitative Finance and Risk Analysis

Working Paper Series Nº 20 - 6:

Likelihood-based Dynamic Asset Pricing: Learning Time-varying Risk Premia from Cross-Sectional Models

Dennis Umlandt

Likelihood-based Dynamic Asset Pricing: Learning Time-varying Risk Premia from Cross-Sectional Models*

Dennis Umlandt†

August 3, 2020

Abstract

This paper proposes a new parametric approach to estimate linear factor pricing models with time-varying risk premia. In contrast to recent contributions to the literature, the framework presented abstains from introducing instrument variables to describe the time variation of risk prices. Instead, time-varying risk prices and exposures follow a recursive updating scheme constructed to reduce the one-step ahead prediction error from a cross-sectional factor model at the current observation. This agnostic approach is particularly useful in situations where instrument variables are unavailable or of poor quality. Estimation and inference are done by likelihood maximization. A Monte Carlo study compares the ability of the method to predict risk prices and returns to that of a regression-based method that uses noisy signals from true risk price predictors. In a realistic setting, the two approaches keep pace when the signal contains 80 percent correct information. An application to a macro-finance model of currency carry trades illustrates the novel approach.

Keywords: Dynamic Asset Pricing, Generalized Autoregressive Score Models, Time-varying Risk Premia, Return Predictability

JEL Codes: G12, G17, C58

*The author thanks Emanuel Moench, Christian Bauer, Alexander Gretener, Markus Haas, Alexander Klos, Christoph Meinerding and Stefan Reitz, as well as participants of the Brown-Bag-Workshop at Kiel University (2019), and the Doctoral Workshop at the 26th Annual Meeting of the German Finance Association (2019) for valuable discussions and comments. Hospitality of Deutsche Bundesbank, where substantial work on this paper has been conducted during a stay in the research center, is gratefully acknowledged.

†University of Trier, Universitaetsring 15, D-54296 Trier, Germany. Email: umlandt@uni-trier.de

1 Introduction

Risk premia for holding financial assets vary over time (Campbell and Shiller (1988), Fama and French (1989), Lettau and Ludvigson (2001), Cochrane (2011)). Traditional factor asset pricing models (for example, Fama and French (1993) and Carhart (1997)) describe these premia with risk prices (lambdas) demanded by investors for each unit of exposure (beta) to a financial or macroeconomic source of risk. Although the predominant methods to test asset pricing models like the two-step regression procedure of Fama and MacBeth (1973) (in the following referred to as the FMB procedure) rely on constant lambdas and betas, the appropriateness of this assumption has been largely doubted (Jagannathan and Wang (1996), Ghysels (1998)).

Recent estimation approaches (Adrian et al. (2015), Gagliardini et al. (2016), Adrian et al. (2019)) formulate lambdas as functions of instrument variables generating the risk price dynamics. These approaches are advantageous for testing asset pricing theories which suggest drivers of risk premia. However, employing inappropriate explaining instruments or leaving out relevant ones may yield to misleading results if one is mainly interested in filtering time-varying risk premia. Empirical research in dynamic asset pricing would therefore benefit from methods that allow for exploring risk price and exposure dynamics implied solely by the cross-sectional model specification rather than dynamics prescribed by additional external forecast variables. Knowledge of these agnostic dynamics can in turn offer a better understanding of risk premia dynamics in asset pricing theories. In addition, financial professionals can benefit from adequately predicted risk premia time series in their hedging and forecasting duties.

The method presented in this paper allows for estimating dynamic versions of cross-sectional factor asset pricing models without specifying forecast or instrument variables for driving time dynamics. In every period, the idea is to evaluate current pricing errors with an observation density, implied by the factor model, to disentangle how much mispricing can be attributed to variation in parameters and update them accordingly. For the choice of the direction and intensity of this parameter updating, I follow the general approach by Creal et al. (2013) and Harvey (2013) who propose to update parameters in econometric models towards the direction of the gradient of the log likelihood, the so-called score, evaluated at the current observation. Hence, parameters are pushed towards the direction with the steepest increase in the likelihood function pointed out by the gradient. I therefore model betas and lambdas as unobserved components whose dynamics are driven by the scaled score of the observation density evaluated at the current observation in order to reduce the one-step-ahead prediction error.

This Generalized Autoregressive Score (GAS)[1] framework by Creal et al. (2013) has been applied successfully in numerous applications in time series analysis and financial econometrics.[2] This papers' main contribution is the introduction of a general framework for dynamic asset pricing models with GAS parameter dynamics. The resulting score-driven likelihood-based asset pricing model (SLAPM) works for every linear (cross-sectional) factor model that can be analyzed using the traditional FMB procedure and generates latent risk price and exposure dynamics. Lambda and beta series are estimated from asset returns and cross-sectional risk factor data only. Estimation and inference can be carried out according to the maximum likelihood principle.

Optimal updating schemes for time-varying lambdas and betas with respect to the observational likelihood are derived in case of elliptically distributed asset returns and show an intuitive relation to the FMB estimates. The updating corrects lambdas and betas with respect to local cross-sectional errors that are produced when employing FMB with constant parameters and thus can be regarded as a local FMB correction. Moreover, the updating reduces the impact of extreme observations when considering heavy-tailed distributions. In particular, the filtered lambdas and betas are therefore more robust against outliers in the data.

[1]Also referred to as Score-Driven (SD) model or Dynamic Conditional Score (DCS) model.

[2]See, for example, Harvey and Lange (2017) and Gorgi et al. (2019) for applications in volatility modeling or Oh and Patton (2018) and Bernardi and Catania (2019) for systemic risk applications.

A Monte Carlo study investigates the ability of a Gaussian SLAPM with constant betas to filter risk price dynamics in a setting with realistic signal-to-noise ratio. The performance of the SLAPM is compared to that of a dynamic asset pricing model developed by Adrian et al. (2015) (in the following referred to as DAPM) which employs noisy signals of different strengths of the true instrument factor driving the lambda dynamics of the data-generating process. In an asset return panel of 25 assets and 500 time observations, a DAPM would need to be informed with a signal containing more than 80% of the information from the true data-generating process to at least be able to compete with a SLAPM as measured by a Diebold and Mariano (1995) test.

Moreover, the simulation study is extended to asset-specific return predictions. Predictive regressions as benchmarks are not able to systematically outperform the SLAPM return forecasts in simulated panels even if they are informed with the correct predictors. The SLAPM can therefore compensate its disadvantage from missing predictor information by filtering cross-sectional return data through the lens of a cross-sectional factor pricing model. Results from the simulation study therefore supports the view that a substantial portion of systematic time series predictability, that is usually studied in a univariate context, should be inferable from cross-sectional pricing errors if the cross-sectional factor model is correctly specified.

An empirical application to the international macro-finance model of Lustig et al. (2011) illustrates the SLAPM's ability to filter time-varying currency premia. Opie and Riddiough (2020) propose a method for gobal currency hedging that exploits predictable components in the two common risk factors of Lustig et al. (2011) estimated with economically motivated forecast factors. A comparison with SLAPM-implied risk premia suggests that currency premium forecast factors capture most risk price movements. However, sudden downfalls in carry trade risk premia that fall in line with currency crashes, that are suspected as potential cause of carry trade risk premia due to Brunnermeier et al. (2008), disconnected from macroeconomic conditions are captured by the SLAPM only. Although these currency crashes are not predictable from economic conditions, they matter for in-sample filtering of currency risk premia.

The DAPM of Adrian et al. (2015) is considered as main benchmark for comparisons. Both the DAPM and SLAPM rely on the same beta pricing equation but differ in the construction of risk price dynamics (external forecasters vs. recursive observation-driven updating) and estimation methodology (three-step linear regressions vs. maximum likelihood). Adrian et al. (2019) employ an enhanced version of the DAPM approach that allows for non-linear relations between risk prices and forecast variables. Risk prices and exposures being affine-linear transformations of instruments are also employed in Gagliardini et al. (2016), Chaieb et al. (2018) and Gagliardini et al. (2019) with a focus on large unbalanced panels of individual stock returns. The functioning and intentions of these methods are different to mine. Their aim is to filter economic meaningful variation in risk prices explained by forecasting factors whereas this work investigates the portion of risk price movement that can be learned from observed cross-sectional model pricing errors only.

Early contributions already allow for instrument-free dynamics of risk prices by conducting cross-sectional FMB regressions period by period as in Fama and MacBeth (1973) and Ferson and Harvey (1991). The approach presented in the following differs from the traditional one by explicitly modelling an intertemporal relation between lambdas of different time periods that can be fitted and analyzed whereas period-by-period FMB risk price estimates are not explicitly connected over time and extremely volatile. The risk price updating mechanism in my approach can be understood as an attempt to infer how much of this risk price volatility stems from actual parameter movements.

The paper is organized as follows. Section 2 introduces and discusses the proposed score-driven dynamic asset pricing framework and closes by laying out a strategy for likelihood-based estimation and inference. A Monte-Carlo study evaluating the performance of the SLAPM is conducted in Section 3. The empirical application to a cross-sectional currency model is presented in Section 4. Section 5 concludes.

2 A Score-driven Likelihood Asset Pricing Framework

The following chapter introduces a framework for score-driven likelihood-based asset pricing models. The model setup is described first and is followed by a derivation for optimal parameter-updating schemes in case of elliptical distributions. Before turning to the applications in the following section, the estimation strategy is explained and discussed.

2.1 Model Setup

The basic model setup outlined in the following is in line with the approach presented in Adrian et al. (2015) but differs in the specification of risk price and exposure dynamics that are driven by scores of the observation density instead of forecasting variables.

Let $r_t = (r_t^1, \dots, r_t^N)^\top$ denote the N-dimensional vector representing the excess returns of N different assets at time $t \in \{0, \dots, T\}$. Let the underlying data-generating processes be defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with a filtration $\mathcal{F}_t = \sigma(\{r_t, \dots, r_0\})$ representing the set of information available at time t. Suppose the risk in the economy is described in terms of K risk factors covered in the state vector f_t that follows an adapted $VAR(1)$ process given by

$$f_{t+1} = \mu + \Phi f_t + u_{t+1}, \qquad t = 0, \dots, T-1. \tag{1}$$

with u_t being i.i.d. disturbances. The modelling approach is not restricted to this choice and more general models like VAR(p) processes may be taken into consideration to better capture the distribution of f_t as this is particularly important for maximum likelihood (ML) estimation. I decide to stick to the VAR(1) model for the sake of brevity and because factors are mostly returns with if-at-all first order auto-correlation (Fama and French (1993), Carhart (1997)).

Assume the existence of a unique stochastic discount factor (SDF) m_t that prices every asset $i \in \{1, \dots N\}$ according to

$$\mathbb{E}_t(m_{t+1} r^i_{t+1}) = 0. \tag{2}$$

The Euler equation (2) can be reformulated to

$$Cov_t(m_{t+1}, r^i_{t+1}) = -\mathbb{E}_t(m_{t+1})\mathbb{E}_t(r^i_{t+1}) \tag{3}$$

with $\mathbb{E}_t$ and Cov_t denoting the conditional expectation and covariance with respect to time t information $\mathcal{F}_t$. Let the $K \times K$ matrix $\Sigma_{u,t}$ be the possibly time-varying covariance matrix of the risk factor innovations u_t. Regressing the demeaned return of asset i on the factor innovations u_{t+1} yields an idiosyncratic noise term $e_{i,t+1}$ that is orthogonal to u_{t+1}. Taken together with (3) the return can be decomposed to

$$r^i_{t+1} = \mathbb{E}_t(r^i_{t+1}) + (r^i_{t+1} - \mathbb{E}_t(r^i_{t+1})) \tag{4}$$

$$= -\frac{Cov_t(m_{t+1}, r^i_{t+1})}{\mathbb{E}_t(m_{t+1})} + \beta_i^\top u_{t+1} + e_{i,t+1} \tag{5}$$

with $\beta_{i,t} = \Sigma^{-1}_{u,t} Cov_t(u_{t+1}, r^i_{t+1})$ being the K-dimensional vector of risk exposures. Let the SDF be affine-linear in the economy's risk factor innovations i.e.

$$\frac{m_{t+1} - \mathbb{E}_t(m_{t+1})}{\mathbb{E}_t(m_{t+1})} = -\lambda_t^\top \Sigma^{-1}_{u,t} u_{t+1} \tag{6}$$

with time-invariant price of risk vector λ_t of dimension K. Plugging the SDF into the return decomposition (5) yields a standard beta representation given by

$$r^i_{t+1} = \lambda_t^\top \Sigma^{-1}_{u,t} Cov_t(u_{t+1}, r^i_{t+1}) + \beta_i^\top u_{t+1} + e_{i,t+1} \tag{7}$$

$$= \beta_{i,t}^\top \lambda_t + \beta_{i,t}^\top u_{t+1} + e_{i,t+1}. \tag{8}$$

The decomposition (8) therefore consists of a predictable risk premium $\beta_{i,t}^\top \lambda_t$ that prices risk exposures and another unpredictable component $\beta_{i,t}^\top u_{t+1}$ depending on risk factor innovations. Representation (8) may be interpreted as a system of seemingly unrelated regressions (SUR) with time-varying coefficients and identical regressors that can be stacked to

$$r_{t+1} = \beta_t(\lambda_t + u_{t+1}) + e_{t+1} \tag{9}$$

with $\beta_t = (\beta_{1,t}, \dots, \beta_{N,t})^\top$. What essentially distinguishes the approach proposed here from prior contributions is the specification of the dynamics of the time-varying $(N+1)K$-dimensional[3] parameter vector $\theta_t = (\lambda_t^\top, vec(\beta_t)^\top)^\top$. Whereas for example Adrian et al. (2015) assume that λ_t is affine-linear in a forecasting variable that has to be specified and time-varying betas are derived from a non-parametric kernel-based approach, the approach discussed here will suggest both sets of varying parameters to be driven by recent observations from the return panel and cross-sectional pricing factors f_t.

The GAS model[4] proposed by Creal et al. (2013) provides an opportunity to introduce time variation into general models with specified observation densities. The basic idea is to let the time-varying parameters of a model be updated proportionally to the score of the observation density i.e. the derivative of the logarithmic density with respect to the parameter that should become time-varying. Thus, the parameter vector is pushed in the direction indicated by the gradient. This is the direction in which the update would yield the steepest increase in the observation density. The approach can therefore be understood as a parameter update optimizing the local likelihood in period t.

In the framework described above, asset returns r_t as well as risk factor realizations f_t are observed and the conditional observation density $p(r_t, f_t|\mathcal{F}_{t-1}, \theta_{t-1})$ has to be specified. Given such a specification, the GAS updating scheme for the dynamic vector of risk prices

[3] K lambdas and N betas per lambda.

[4] Models of this type are also known as dynamic conditional score (DCS) or score-driven models. See Harvey (2013).

and exposures is then given by:

$$\theta_t = \omega + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_j \theta_{t-j} \tag{10}$$

$$s_t = \mathcal{I}_t^{-1} \nabla_t, \qquad \nabla_t = \frac{\partial \ln p(r_t, f_t | \mathcal{F}_{t-1}, \theta_{t-1})}{\partial \theta_{t-1}}. \tag{11}$$

Equation (10) determines the updating mechanism for the time-varying parameter θ_t. Here, ω is an $(N+1)K$-dimensional vector of intercepts and A_i, B_j are $(N+1)K \times (N+1)K$ matrices of coefficients for $i = 1, \dots, p$ and $j = 1, \dots, q$. The updating process therefore consists of a constant part, an adjustment of the observation density score, and an autoregressive part.

The centerpiece of score-driven models is the specification of the innovation sequence s_t. This is done by setting s_t proportional to the so-called score ∇_t defined in (11). Creal et al. (2013) leave open to scale the impact of the score sequence to make models more robust to outliers. I follow their proposal to employ the inverted Fisher information matrix $\mathcal{I}_t^{-1}$ for scaling. The specific definition in our framework is $\mathcal{I}_t = \mathbb{E}_{t-1}(\nabla_t \nabla_t')$ with $\mathbb{E}_{t-1}$ being the expectation operator with respect to $p(r_t, f_t | \mathcal{F}_{t-1}, \theta_{t-1})$. This bears the advantage that the scaling depends directly on the variance of the score and yields an intuitive interpretation of the updating scheme if elliptically distributed residuals are assumed as will be shown below.[5] Given that $\mathbb{E}(\nabla_t \nabla_t')$ is finite, the innovation s_t is a martingale difference sequence implying that θ_t in (10) is an ARMA process that inherits the features of this class of time series. This holds particularly for stationarity conditions.

A possible alternative to specify score-driven risk parameters would be to model s_t in (10) as i.i.d. innovation. This would head towards a variant of the popular Kalman filter. However, remark the crucial difference that in a Kalman filter setting s_t would be another independent source of randomness whereas the mechanism in (11) just relates s_t to the innovations e_t and u_t that are already existent in the model. The mechanism therefore captures systematic variations in idiosyncratic and factor innovations to generate movements

[5] Other popular scalings use an identity matrix or the Cholesky factor of the inverse Fisher information. The latter choice may be attractive to achieve unit variance of the scaled score process s_t.

in risk exposures and prices. This will become clearer when specifying distributions in the following.

2.2 Optimal Parameter Updating for Elliptically Distributed Returns and Factors

To compute the score ∇_t, assumptions on the distributions of the innovations u_t and e_t have to be made. I consider the distribution families from the general class of elliptical distributions[6] to allow for flexibility in fitting the data while keeping the framework tractable. The general N-dimensional density of such distributions can be formulated as $p(x) = |\Omega|^{-\frac{1}{2}}\psi\left(x'\Omega^{-1}x, \nu\right)$ with a dispersion matrix $\Omega \in \mathbb{R}^{N\times N}$, a degrees-of-freedom parameter $\nu \in \mathbb{R}_{>0}$, and a characteristic generator $\psi : \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$[7] . The elliptical class includes not only the normal distribution, as can be seen by choosing the generator to be $\psi(x,\nu) = (2\pi)^{-N/2}\exp\left(-x/2\right)$, but also many other distributions like the Laplace and the Student's t-distribution that are popular for fitting financial returns.

Quite intuitive formulas can be derived for the driving martingale difference sequences s_t of time-varying risk prices and exposures can be derived when assuming zero cross-information quantities i.e. $\mathcal{I}_t^{\lambda,\beta} = \mathcal{I}_t^{\beta,\lambda} = 0$. This assumption additionally assures the validity of the derived updating schemes below, when betas or lambdas are assumed to be constant what often yields more parsimonious models. Explicit calculations can be found in Appendix A. The scaled scores turn out to be

$$\begin{pmatrix} s_t^{\lambda} \\ s_t^{\beta} \end{pmatrix} = C(\tilde{e}_t, \psi^e)\begin{pmatrix} (\beta_{t-1}^{\top}\Omega_e^{-1}\beta_{t-1})^{-1}\beta_{t-1}^{\top}\Omega_e^{-1}r_t - (\lambda_{t-1}+u_t) \\ vec(e_t(\lambda_{t-1}+u_t)^{\top}(\lambda_{t-1}\lambda_{t-1}^{\top} + \mathbb{E}_{t-1}(u_t u_t^{\top}))) \end{pmatrix} \tag{12}$$

[6]See Fang et al. (1990) or Chapter 6 of Embrechts et al. (2015) for a comprehensive treatment of elliptical distributions.

[7]The general definition of elliptical distributions includes a location parameter vector $\mu \in \mathbb{R}^N$. However, a zero mean of the innovation terms is assumed throughout the paper and μ is therefore neglected for convenience.

with scalar factor

$$C(\tilde{e}_t, \psi^e) = -\frac{N\frac{\psi_1^e}{\psi^e}}{2\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1^e}{\psi^e}\right)^2\right)}. \tag{13}$$

where $\|\cdot\|$ refers to the Euclidean norm and $\tilde{e}_t = \Omega_e^{-1/2} e_t$ is the standardized idiosyncratic residual.

The updating sequence (10) in conjunction with the scaled scores in (12) can be regarded as a scaled difference of the generalized least squares (GLS) coefficient estimate from cross-sectionally regressing r_t on β and $\lambda_{t-1} + u_t$. The driving mechanism therefore intuitively corrects local deviations in the cross-sectional fit. This observation can be related to the very popular estimation approach of Fama and MacBeth (1973) in which constant risk prices are estimated with a cross-sectional regression of the average portfolio returns $\overline{r}$ on their betas i.e. the risk price estimate is given by $(\beta^\top \Omega_e^{-1}\beta)^{-1}\beta^\top \Omega_e^{-1}\overline{r}$. The correction step proposed by (12) for the lambda essentially enforces this average prescription. Finding $\lambda_{t-1} + u_t$ to be greater than $(\beta_{t-1}^\top \Omega_e^{-1}\beta_{t-1})^{-1}\beta_{t-1}^\top \Omega_e^{-1} r_t$ would be interpreted as an overestimation of the market price of risk plus factor innovation. The resulting s_t^λ would be negative (given that $C(\tilde{e}_t, \psi^e) > 0$) and would downsize the market price of risk for the next period. The term $C(\tilde{e}_t, \psi^e)$ as well as the coefficient matrices $A_1, \ldots A_p$ reveal how much of the local estimation error can be possibly attributed to a change in risk prices and not to factor or idiosyncratic innovations. This mechanism points out the difference of the proposed method to the time-varying lambda framework in Fama and MacBeth (1973) which would choose λ_t in each period to minimize the cross-sectional regression error. This comes with the drawback that time-varying lambdas become unrealistically volatile because cross-sections are fitted independently period by period and do not draw information from connections between lambdas of different time periods. The optimal beta updating analogously adjusts to the local OLS error from regressing r_t on $\lambda_{t-1} + u_t$ as prescribed by the stacked SUR model in (9) while assuming a stochastic regressor u_t[8].

Also remark the similarities to the popular generalized conditional heteroscedastictiy (GARCH) time series models by Engle (1982) and Bollerslev (1986). If the time-varying parameter in (10) would be the conditional variance of a zero-mean time series ε_t, i.e. $\theta_t = \sigma_t^2$ and the score be the squared current observations, i.e. $s_t = \varepsilon_t^2$, the famous GARCH(p,q) updating equation would be obtained.[9] This updating is quite intuitive since in a setting with a constant variance parameter, it would be estimated via the mean of squared observations. The statistic of new incoming information ε_t^2 in relation to past observation therefore tells whether it is likely that the variance has increased (if ε_t^2 is relatively high) or has decreased vice versa. SLAPM works with the same principle as parameters are increased/decreased according to the cross-sectional pricing errors after adjusting for risk factor innovations.

With regard to the distributional assumption it is striking that the particular choice of the elliptical return distribution does not alter the direction of the score and therefore also does not alter the direction of the parameter updating. However, the scalar part depends on the common distribution of asset-specific residuals represented by ψ^e. Updating steps are therefore weighted with regard to the shape of the particular distribution.

A particularly simple updating scheme is achieved when assuming normally distributed innovations, as $C(\tilde{e}_t, \psi^e) \equiv 1$ holds in this case.[10] The effect of the scaling becomes clear when looking at the Student's t-updating sequence, where the scalar part is given by $C(\tilde{e}_t, \psi^e) = \frac{\nu+N+2}{\nu+e_t'\Omega_e^{-1}e_t}$.[11] The division by $\|\tilde{e}_t\|^2 = e_t'\Omega_e^{-1}e_t$ down-weighs the impact of scores when observing extreme values and therefore takes the more pronounced tails of the return distribution into account. Hence, imposing this updating structure makes the market price of risk more robust to outliers in the data. Therefore, it might henceforth be beneficial for tractability to assume a distribution that yields a simple updating scheme such as the

[8]Note that f_t is actually assumed to be observed and not u_t. Assuming u_t being an observed factor itself would analogously yield an updating according to the corresponding error from regressing r_t on $\lambda_{t-1} + u_t$ while assuming deterministic u_t.

[9]Creal et al. (2013) argue that the Gaussian GARCH(1,1) model is indeed a GAS-model.

[10]See Corollary 3 and its proof in the technical appendix.

[11]See Corollary 4 and its proof in the technical appendix.

Gaussian distribution. However, such a simplification would come at the cost of incorrectly weighting the magnitude of extreme observations.

Moreover, only the updating scheme for risk exposures β_t does depend on the distribution of the factor innovations u_t (via $\mathbb{E}_{t-1}(u_t u_t^\top)$). Optimal risk price updating is not altered at all by different factor model specifications. Hence, in a constant beta setting, one could easily introduce some statistical features like heavy tails and asymmetries in the model without giving up a simple updating scheme by adjusting the distribution of u_t. However, this does not open up the possibility to allow these features idiosyncratically for specific assets in the cross-section as this would come with the necessity to leave the class of elliptical distributions.

With the specified distributions for the residuals, (12) completes the dynamic asset pricing model framework consisting of the equations (1), (9), (10), and (12). I refer to it as the (elliptical) SLAPM of orders p and q (SLAPM(p,q)).

2.3 Estimation and Inference

The elliptical SLAPM discussed so far ends up with a reduced-form representation that can be estimated with an ML-procedure. The set of static parameter vectors and matrices to be estimated is given by $\omega, A_1, \dots A_p, B_1, \dots B_q, \mu, \Phi, \Omega_e, \Omega_c, \nu_e$, and ν_u. This collection of parameters needs to be chosen to maximize the log-likelihood function given by

$$\mathcal{L} = \sum_{t=1}^{T} \ln \psi^e(e_t' \Omega_e^{-1} e_t, \nu_e) + \ln \sum_{t=1}^{T} \psi^u(u_t' \Omega_u^{-1} u_t, \nu_u) - \frac{T}{2} \left(\ln |\Omega_e| + \ln |\Omega_u| \right). \tag{14}$$

Besides the problems associated with the enormous number of parameters, closed-form solutions of the ML-estimators are not available due to the strong dependencies of the parameters on each other. This makes it necessary to employ numerical optimization procedures.

The elliptical SLAPM(p,q) is quite general and can be further simplified by imposing technical or economic restrictions, which would result in a more parsimonious model. As an illustration of the framework, let us focus on a Gaussian SLAPM with constant betas which yields a reduced number of parameters. Although there is a large body of literature assuming time-varying betas, empirical studies such as Braun et al. (1995) and Ghysels (1998) document that changes in betas are rather slow-moving. Assuming constant betas and letting the dynamics kick in from time-varying lambdas to achieve a parsimonious model could therefore be justified in this regard. The Gaussianity of the asset's residual distribution represented by the characteristic generator ψ^e ensures a simple and tractable updating scheme for the time-varying lambda that adjusts for local Fama-MacBeth errors. These assumptions turn out to be sufficient to reasonably filter risk price dynamics from asset as the following simulation study and the empirical application document. However, for prediction purposes, a more adequate fit to the statistical properties of return time series can be achieved by an alternative specification of the factor innovation distributions u_t as mentioned in the previous section.

The complete reduced form Gaussian constant beta SLAPM(p,q) is then given by

$$f_{t+1} = \mu + \Phi f_t + u_{t+1}, \qquad t = 0, \ldots, T-1 \tag{15}$$

$$r_{t+1} = \beta(\lambda_t + u_{t+1}) + e_{t+1} \tag{16}$$

$$\lambda_t = \omega^\lambda + \sum_{i=1}^{p} A_i^\lambda s_{t-i+1}^\lambda + \sum_{j=1}^{q} B_j^\lambda \lambda_{t-j} \tag{17}$$

$$s_t^\lambda = (\beta^\top \Omega_e^{-1} \beta)^{-1} \beta^\top \Omega_e^{-1} r_t - (\lambda_{t-1} + u_t) \tag{18}$$

The likelihood in (14) can be optimized in several steps. The parameters of the factor VAR model (15) are estimated independently in a first stage where the ML-estimator is the ordinary least squares estimator in the Gaussian case. This estimation approach may yield inefficiencies because the factor innovations u_t enter the idiosyncratic portfolio errors e_t via the return equation (16). However, this effect is found to be negligible.

In the second step, the fitted residuals from (15) are plugged in and the likelihood is numerically optimized with respect to the remaining GAS parameters $\lambda_0, \omega^\lambda, A_1^\lambda, \ldots A_p^\lambda$,

$B_1^\lambda, \ldots, B_p^\lambda$, the dispersion matrix Σ_e and risk exposures β using a quasi-Newton procedure. The number of parameters can be further reduced when assuming cross-sectionally homoscedastic errors e_t for deriving the updating scheme in (18) that henceforth simplifies to $s_t^\lambda = (\beta^\top \beta)^{-1}\beta^\top r_t - (\lambda_{t-1} + u_t)$. Because of the missing occurrence of Ω_e in the updating scheme, the maximum likelihood estimator of the dispersion matrix is given by $\hat{\Omega}_e = \frac{1}{T}\sum_{t=1}^{T} e_t e_t^\top$. This approach may weight innovations in the updating scheme incorrectly, in particular if idiosyncratic errors are highly correlated. However, one can construct a correction by running a second maximum likelihood estimation with a pre-specified dispersion matrix estimated from a first estimation stage with homoscedastic errors in the spirit of feasible GLS estimation approaches.

Inference is conducted in the standard fashion for ML-estimators as suggested by Creal et al. (2013) for GAS models in general. General conditions for consistency and asymptotic normality for general GAS models are provided in Blasques et al. (2014). If ϑ stacks all the static parameters of the model, standard asymptotic theory for ML-estimators would suggest that under some regularity conditions the following holds:

$$\sqrt{T}\left(\hat{\vartheta} - \vartheta\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{I}^{-1}(\vartheta)\right) \tag{19}$$

with Fisher information matrix $\mathcal{I}(\vartheta) := -\mathbb{E}\left(\partial^2 l_t / \partial\vartheta\partial\vartheta^\top\right)$ where l_t is the log-likelihood contribution of the i-th observation evaluated at ϑ. Since the use of a Gaussian SLAPM is a potential source for model mis-specification, it could be advisable to rely on quasi ML standard errors that can be derived with

$$\sqrt{T}\left(\hat{\vartheta} - \vartheta\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{I}^{-1}(\vartheta)\mathcal{J}(\vartheta)\mathcal{I}^{-1}(\vartheta)\right) \tag{20}$$

where $\mathcal{J}(\vartheta) := \lim_{T\to\infty} T^{-1}\mathbb{E}\left(\left(\partial\mathcal{L}(\vartheta)/\partial\vartheta\right)\left(\partial\mathcal{L}(\vartheta)/\partial\vartheta\right)^\top\right)$. Standard errors in the following are derived by numerically differentiating the score function with a finite difference approximation.

3 Monte Carlo Study

The performance of the Gaussian constant beta SLAPM is evaluated with a Monte Carlo study on forecasting risk prices and excess returns in the following. A main question to answer is whether the SLAPM is able to filter risk price movements without knowing the driving forces and how well it can compete with models making use of information about the predictable component.

3.1 Data-Generating Process

The DAPM of Adrian et al. (2015) is considered as main benchmark method for comparison. It assumes that factors in f_t may be risk factors f_{1t} or[12] forecasting factors f_{2t} in such a way that $f_t = (f_{1t}^\top, f_{2t}^\top)^\top$ and

$$r_{t+1} = \beta(\lambda_0 + \Lambda_1 f_{2t}) + \beta u_{1,t+1} + e_{t+1} \tag{21}$$

with $u_{1,t+1}$ being the innovations to the risk factors f_{1t} from the factor VAR model (1). The DAPM can be estimated with a three-step linear regression approach and nests the constant risk price model of Fama and MacBeth (1973).

The data-generating process employed for simulations is chosen in line with the DAPM modeling approach with exactly one risk factor f_{1t} and one forecasting factor f_{2t}. This poses a challenge for the SLAPM procedure that has to prove itself within the framework of a competing model approach. In order to generate realistic returns we orientate on a capital asset pricing model (CAPM) of an industrial portfolio cross-section from the Kenneth French data library. The pricing factor supposed to be the market return that is found to be fairly well represented by an ARMA(1,1)-process with GARCH(1,1)-residuals. Thus, the heteroscedasticity in the return series is modeled to be sourced from the risk factor process.

[12] Adrian et al. (2015) explicitly include the possibility for factors to be both risk and forecasting factors. For the simulation study, I abstain from this for simplicity.

To get a candidate process for the forecasting factor, a DAPM with two forecasting factors is fitted. The forecasting factors under consideration are the 10-year treasury yield obtained from the H.15 statistical release of the Board of Governors of the Federal Reserve System as well as the dividend yield of the $S\&P$-500 index. Both series are found to be reasonable stock return predictors in the past.[13] For simplicity, the two predictors are combined to form one forecast process by computing the linear combination of the two weighted by coefficients from the fitted DAPM. The forecasting factor process dynamics are found to be well-described by an AR(1)-process given by $f_{2,t+1} = 0.005 + 0.98 f_{2,t} + u_{2,t}$, $u_{2,t} \sim \mathcal{N}\left(0, \sigma_f^2\right)$ and $\sigma_f = 0.11$ that is used for simulation in the following. The high level of persistence with AR(1)-coefficients near unity is often observed for return forecasting factors (Campbell and Yogo (2006)) are in line with long-run risk models like Bansal and Yaron (2004) featuring small but persistent predictable components. From the simulated risk and forecasting factor processes, the risk price process is derived as $\lambda_t = 0.38 + f_{2,t}$, where the intercept has been calibrated to the industry portfolio data. The simulated return values are finally derived from the beta representation according to (21). The first three betas take the values $0.5, 1$ and 1.5 while the remaining are uniformly drawn from the interval $[0.5, 1.5]$ that is roughly the range observed for portfolio exposures when regressing portfolio returns on the market risk factor.

[Figure 1 about here.]

Figure 3(a) shows an example draw of the simulated excess return of one portfolio from cross-section ($N = 25, T = 500$) together with its conditional excess return $\beta\lambda_{t-1}$. The conditional return in relation to the actual return can be regarded as only fairly small in magnitude. This fits the observation that stock returns show only little predictability, if at all.

[13] See Keim and Stambaugh (1986), Campbell (1987), Fama and French (1989), Campbell and Thompson (2008) for evidence on long-run treasury yields and Campbell and Shiller (1988), Fama and French (1989), Campbell and Thompson (2008), Cochrane (2008) for dividend yields.

3.2 Predicting Risk Prices: Scores versus Forecasters

Because $\mathbb{E}_T(r_{T+1}) = \beta\lambda_T$, predicting stock market returns for period $T+1$ would need an accurate estimation of λ_T. The performance of the SLAPM is therefore evaluated by its ability to predict the one-period-ahead risk price in comparison to the corresponding ability of the DAPM and FMB regressions. Predictions implied by the different approaches are:

$$\text{SLAPM}: \quad \hat{\lambda}_T^{SLAPM} = \hat{\omega} + \hat{A}s_T + \hat{B}\hat{\lambda}_{T-1}^{SLAPM} \tag{22}$$

$$\text{FMB}: \quad \hat{\lambda}_T^{FMB} = \hat{\lambda}^{FMB} \tag{23}$$

$$\text{DAPM}: \quad \hat{\lambda}_T^{DAPM} = \hat{\lambda}_0 + \hat{\Lambda}_1 f_{2,T}. \tag{24}$$

Given the simulated data set, λ_T is predicted with the three approaches mentioned above. Since the part of the data-generating process concerning the market price of risk dynamics completely follows the specification in Adrian et al. (2015), the DAPM estimator has a trivially high information advantage over the other two approaches when using $f_{2,t}$. To circumvent this issue, the DAPM is fitted with employing a diffuse signal of the true forecast factor realization ranging from pure noise to the true signal. The DAPM is therefore estimated with a signal $\tilde{f}_{2,t}^{\kappa}$ drawn from

$$\tilde{f}_{2,t}^{\kappa} = \kappa f_{2,t} + (1-\kappa)\epsilon_t, \qquad \epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_f^2) \tag{25}$$

with $\kappa \in [0,1]$. For $\kappa = 0$ the signal exhibits no information about the true factors and for $\kappa = 1$ the DAPM exploits the true forecasting factor series. The results will of course heavily depend on the variance of the noise term ϵ_t. I decide to set this variance parameter to be equal to the variance of the true predictor process f_t for improved comparability.

Figure 3(b) exemplary shows that the SLAPM seems to be quite effective in filtering λ_t. The updating direction and turning points are anticipated correctly although with some delay. Only peaks are hard to capture for the likelihood score-based filter. This means that extreme local cross-sectional errors resulting due to extreme changes in the risk price are devoted with a too large share to factor and idiosyncratic innovations. The full-information DAPM, as expected, captures the peaks slightly but has a tendency to overshoot.

For achieving a reliable evaluation of the SLAPM's risk price forecast ability, several N-T-panels with $N \in \{25, 50, 100\}$ assets and $T \in \{250, 500, 1000\}$ time observations are simulated with $S = 10\ 000$ replications each. The predictive accuracy is evaluated with out-of-sammple R^2 (OOS-R^2) proposed by Campbell and Thompson (2008). This is computed as

$$R^2_{OOS} = 1 - \frac{\sum_{s=1}^{S} \left(\hat{\lambda}_T^{SLAPM}(s) - \lambda_T(s) \right)^2}{\sum_{s=1}^{S} \left(\hat{\lambda}_T^{i}(s) - \lambda_T(s) \right)^2} \tag{26}$$

with benchmark forecast from $i \in \{FMB, DAPM\}$. Hence, whenever the quantity is positive, the SLAPM forecasts the risk price with a lower squared error than the benchmark and vice versa if the quantity is negative. Diebold and Mariano (1995) (DM) tests on the null hypothesis of equal forecast accuracy are conducted as well.

[Table 1 about here.]

Table 1 shows the results of the risk price forecasting exercise with shaded areas indicating that the DM-test does not reject the null of equal forecast accuracy on a five percent significance level. At first sight, the SLAPM always shows a positive OOS-R^2 with respect to forecasts from the Fama-McBeth regressions, and the null hypothesis of the DM test can be rejected indicating superior forecast accuracy with respect to the FMB forecast. The same holds for the DAPM forecasts with pure noise forecasters. The evidence therefore suggests that the SLAPM succeeds in filtering information on the time variation in risk prices from the return series.

The SLAPM appears to perform particularly well in panels with few time series observations as $T = 250$ where it outperforms the DAPM irrespective of the employed signal. In particular, the DAPM forecast is still inferior when employing the true predictor. This changes crucially with more observations in time. The DAPM with κ of 0.9 and 1 cannot be outperformed by the SLAPM in terms of the DM-test in a 25-500 panel whereas the opposite still holds true for lower κ. This kind of "break-even" κ appears to monotonically decrease with the sample length T. However, such a clear pattern is not observed for the cross-sectional dimension N. This indicates that regression-based methods introducing risk price dynamics with instruments or forecasters are particular promising for particular long panels. This also suggests that the problem of inappropriate instruments may not be resolved by solely increasing the cross-sectional dimension. However, if the sample length is already large with T=1000, the DAPM seem to improve compared to the SLAPM with increasing number of test assets N. Practically relevant panels often have around 500 or less time observations. Remarkably, the SLAPM never gets outperformed by the DAPM even with full information in this short to medium length panels. In the case of $T = 500$ that would correspond to more than 40 years of monthly observations, the DAPM would need a κ of at least 0.9 for not getting outperformed by the SLAPM.

The SLAPM therefore succeeds to filter risk price movements by solely evaluating period-by-period cross-sectional asset pricing errors and can compete with the regression-based models employing external information about the correct driving forces. Another baseline is that even moderate noise in the predictor variable can seriously impair the accuracy of the DAPM in forecasting risk prices. This speaks in favor of the SLAPM as promising alternative not only in situations where forecasters are unknown, but also whenever there is doubt about the correct ones. Regression-based approaches appear to be particular accurate in panels with many time observations ($T \geq 1000$) that would, however, require over 84 years of monthly and more than 250 years of quarterly observations.

3.3 Predicting Returns: Scores versus Predictive Regressions

The previous section compares empirical dynamic asset pricing methods by their ability to forecast risk prices. This is motivated by the fact that return forecasts heavily rely an adequately estimated lambda. However, when it comes to forecasting the specific return of asset i, it is also important to adequately estimate the risk exposure β_i as its expected return is given by $\mathbb{E}_T\left(r^i_{T+1}\right) = \beta_i \lambda_T$.

In order to investigate whether the results from the previous section translate into asset-specific return predictability, the Monte Carlo study is repeated for univariate time series. The return series of asset i is generated as described above with

$$
\begin{aligned}
r^i_{t+1} &= \beta_i(\lambda_0 + \Lambda_1 f_{2t}) + \beta_i u_{1,t+1} + e_{i,t+1} && (27)\\
&= \beta_i \lambda_0 + \beta_i \Lambda_1 f_{2t} + \beta_i u_{1,t+1} + e_{i,t+1} && (28)\\
&= \alpha_0 + \alpha_1 f_{2t} + \varepsilon_{i,t+1} && (29)
\end{aligned}
$$

where $\alpha_0 = \beta_i \lambda_0, \alpha_1 = \beta_i \Lambda_1$ and $\varepsilon_{i,t+1} = \beta_i u_{1,t+1} + e_{i,t+1}$. A tough benchmark for the SLAPM to beat would therefore be a predictive regression employing the predictor f_{2t}. This benchmark is employed in the following, again with signals $\tilde{f}^{\kappa}_{2,t}$ of different level of diffusion κ in line with (25). Simple sample average forecasts are additionally considered such that we have following three candidates:

$$
\begin{aligned}
\text{SLAPM :} &\quad \hat{r}^{i,SLAPM}_{T+1} = \hat{\beta}^{SLAPM}_i \hat{\lambda}^{SLAPM}_T && (30)\\
\text{Mean :} &\quad \hat{r}^{i,Mean}_{T+1} = \frac{1}{T}\sum_{t=1}^{T} r^i_t && (31)\\
\text{PReg :} &\quad \hat{r}^{i,PReg}_{T+1} = \hat{\alpha}_0 + \hat{\alpha}_1 \tilde{f}^{\kappa}_{2,T} && (32)
\end{aligned}
$$

where the coefficients $\hat{\alpha}_0$ and $\hat{\alpha}_1$ in PReg are estimated by regressing r^i_t on $\tilde{f}^{\kappa}_{2,t-1}$. Simulations are again conducted for several N-T-panels with $S = 10\ 000$ replications each. The first three assets in the data-generating process have a low ($\beta_1 = 0.5$), medium ($\beta_2 = 1$) and high ($\beta_3 = 1.5$) beta. These assets are included in each N-T-panel and generate the three return series under investigation.

[Table 2 about here.]

Table 2 provides the results from forecasting simulated returns. Be aware that OOS-R^2's are now given as percentages and the individual return mean squared forecast errors of the alternative methods are much closer to each other than it is the case for risk price forecasts in the previous section. The SLAPM forecasts outperform the Mean forecasts as can be seen in the first rows of the panels (a), (b) and (c). The finding from the previous section, that the proposed method is able to extract useful information of risk price movements, therefore also holds for conditional means of asset-specific returns.

Another main observation is that only in one scenario the SLAPM forecasts is outperformed according to the DM-test in the simulated return panels. This is the panel with N=25 assets and T=500 observations when forecasting the high exposure asset return ($\beta = 1.5$) and the alternative forecast is a predictive regression with correct predictor ($\kappa = 1$). Apart from that case, correct predictors do not empower predictive regressions to generate more adequate forecasts that the SLAPM in the other return panels. In particular, the SLAPM even outperforms correctly specified predictive regressions in short panels (T=250). For a κ lower or about 0.5, the SLAPM forecasts are also more adequate than the predictive regressions with diffused predictor signals in longer panels ($T = 500, 100$). This holds for low, medium as well as high beta asset returns.

With respect to the tough benchmark of predictive regressions that are perfectly specified in the given simulation setting, the SLAPM performs considerably well in forecasting asset-specific returns. The results indicate that the SLAPM is able to compensate its informational disadvantage from lacking forecaster data by considering information from the whole cross-section that is not available to predictive regression forecasts. They support the view that if returns have a cross-sectional factor structure as proposed in the majority of asset pricing models, systematic time series predictability, that usually studied in a univariate context, should be inferable from cross-sectional pricing errors.

4 Empirical Application

The Monte Carlo study in Section 3 provides evidence that the SLAPM performs particularly well compared to predictive regressions in cross-sections simulated from correctly specified models. The following empirical application intends to shed light on whether and to which extent the SLAPM can detect predictable components in real world risk factors. I choose to analyze a dynamic version of the international macro-finance model of Lustig et al. (2011) that prices a cross-section of currency carry trades. These are popular speculative currency trading strategies that usually invest money in currencies bearing high interest rates and finance this investment with a credit in currencies with low interest rates. Lustig et al. (2011) find that this cross-section can be priced by two risk factors only. The first factor (DOL) is the return series of a currency market portfolio that holds an equally-weighted position in each currency. This may be seen as an equivalent to the market factor in the CAPM for equity cross-sections. The second factor (HML) is the return of a portfolio that holds a long position in the high interest rate portfolio financed through a short position in the low interest rate portfolio.

I chose this model as test object for the SLAPM for several reasons. Opie and Riddiough (2020) show that each of the two currency risk factors indeed has a predictable component that can potentially be detected. DOL and HML are furthermore actually executed trading strategies for which forecasters have been researched. This facilitates the search for an adequate forecasting benchmark. Moreover, the two currency factors fit the cross-sections of carry trade returns very well. This to some extent secures that the performance of the SLAPM is not crucially affected by cross-sectional model misspecification.

4.1 Data

The data to be consiered for the cross-sectional model and the following portfolio sorting practice is in line with several contributions in international macro-finance that explain cross-sectional variations in currency cross-sections (Lustig et al. (2011), Menkhoff et al. (2012), Mueller et al. (2017)). Data on currency portfolio return predictors is additionally collected for a benchmark estimation of risk premia implied by economic forecast variables.

4.1.1 Currency Data

The currency data set consists of 48 spot exchange rates (s_t^i in logs) and one-month-forward exchange rates (fwd_t^i in logs) from Thomson-Reuters Eikon with the US-dollar being the base currency and covers the period from April 1986 to November 2018. Currency carry trades are typically implemented with forward contracts. I therefore compute continuously compounded returns of individual currency i as the return to a one dollar investment in the corresponding one-month-forward contract, i.e.

$$r_{t+1}^i = fwd_t^i - s_{t+1}^i. \tag{33}$$

Each currency is allocated into one of five equally-weighted portfolios sorted by their forward discounts $fd_t^i = fwd_t^i - s_t^i$.[14] The first portfolio (C1) therefore includes the fifth of currencies with the lowest interest rate differential to the US and the fifth portfolio (C5) the fifth of currencies with the highest interest rate differential. Transaction costs are adjusted by considering bid and ask quotes. The first risk factor of the model from Lustig et al. (2011) is the return series of a currency market portfolio referred to as DOL-portfolio that holds an equally-weighted position in each currency. This may be seen as an equivalent of the market return factor in the CAPM for equity cross-sections. The second risk factor is the return of a portfolio that holds a long position in the high interest rate portfolio C5 financed through a short position in the low interest rate portfolio C1. This second factor is referred to as HML.

[Figure 2 about here.]

[14] This sorting is equivalent to sorting on the interest rate differential if the covered interest parity holds.

Figure 2 shows cumulative log return of both the DOL and the HML portfolio. The profitability of both series seem to vary substantially over time and therefore suggest that risk premia to be filtered are time-varying as well.

4.1.2 Forecasting Factors

I tested several potential currency premium forecasters motivated from the literature[15] with the regression-based dynamic asset pricing approach of Adrian et al. (2015) and include those that show a significant impact on time-varying DOL and HML risk prices in the benchmark dynamic asset pricing model. The included factors are:

Absolute Average Forward Discount. Cary trade returns stem from favorable spot rate changes and/or differentials in interest rates proxied by forward discounts. If the average forward discount is high in magnitude, the potential for high carry trade returns is increased as returns from interest differentials would be realized if unfavorable spot rate changes hold off. Moreover, Lustig et al. (2014) also find that returns to the dollar factor can be predicted with average forward discounts. I construct the absolute average forward discount $|\overline{fd}_t|$ with

$$\overline{fd}_t = \frac{1}{|I_t|} \sum_{i \in I_t} fd_t^i \tag{34}$$

where I_t is the index set of available currencies in month t.

FX volatility. Carry trades tend to perform particularly bad in times of high uncertainty and are negatively related to measures of foreign exchange market volatility (Bhansali (2007), Menkhoff et al. (2012), Bakshi and Panayotov (2013)). As a measure of FX market uncer-

[15] I mainly considered the forecasting factors proposed in Bakshi and Panayotov (2013) and Opie and Riddiough (2020).

tainty I compute global FX volatility in line with Menkhoff et al. (2012), i.e.

$$VOL_t = \frac{1}{T_t} \sum_{\tau \in T_t} \left(\frac{1}{|I_\tau|} \sum_{i \in I_\tau} |r_\tau^i| \right) \tag{35}$$

where r_τ^i is the log-return of currency i, T_t the number of trading days in month t, and I_τ the index set of available currencies on trading day τ.

Commodity returns. Typical commodity currencies frequently show up in the high-yielding-interest portfolio and commodity currency exchange rate predictability from associated commodity price returns is documented in several studies (Chen and Rogoff (2003), Rossi (2013)). Commodity prices are proxied by the industrials subindex of the CRB spot commodity index (CRB_t). Bakshi and Panayotov (2013) provide regression evidence for carry trade predictability of returns to the CRB index.

Intermediary capital. Adrian et al. (2011) and Brunnermeier and Pedersen (2009) stress the importance of the funding conditions of financial intermediaries for currency risk premia. I consider the aggregated capital ratio (Cr_t) of financial intermediaries referred to as the primary dealers by the New York Fed taken from He et al. (2017).

Dollar competitiveness. The broad dollar index of the Federal Reserve Board (FRB_t) is considered as well. This index is a trade-weighted average of nominal dollar exchange rates against currencies of a broad group of major US trading partners. A stronger dollar is suspected to represent the shadow costs of bank balance sheet capacity (Avdjiev et al. (2019)) and could therefore proxy an impact on currency returns via a bank risk-taking channel.

Following Bakshi and Panayotov (2013) and Opie and Riddiough (2020), I compute for all forecasting factors except the absolute average forward discount 3-month the average log growth rates, i.e.

$$\Delta F_t = \frac{1}{3} \log \left(\frac{F_t}{F_{t-1}} \right) \tag{36}$$

for forecasting factor F_t.

4.2 Empirical Results

The pricing equation in line with our baseline framework of Section 2 is given by

$$r^i_{t+1} = \beta_i^{DOL}(\lambda_t^{DOL} + u_{t+1}^{DOL}) + \beta_i^{HML}(\lambda_t^{HML} + u_{t+1}^{HML}) + e_{i,t+1} \tag{37}$$

In order to introduce time dynamics in risk prices, a Gaussian constant beta SLAPM(1,1) is employed with non-diagonal elements in the parameter matrices A_1^λ and B_1^λ from the updating scheme (17) being restricted to zero.[16] In this way a more parsimonious model is obtained with two risk price updating equations given by

$$\text{SLAPM}: \qquad \lambda_t^j = \omega_j + a_j s_{j,t} + b_j \lambda_{t-1}^j \tag{38}$$

with scalar parameters ω_j, a_j and b_j with $j = DOL, HML$. Two established benchmark specifications are considered to investigate whether the SLAPM approach can actually improve on filtering risk premia and explain return variations in cross-section and time simultaneously. The first benchmark is the constant risk price specification underlying classical Fama and MacBeth (1973) regressions. Estimates and standard errors for this specification are achieved by estimating a DAPM according to Adrian et al. (2015) without forecast variables that nests the FMB estimators but adjust for cross-asset correlation in the residuals. The second benchmark is a DAPM that explains risk price variations with the forecasting factors

[16] The diagonalization mutes the impact of scaled scores and and factor innovation on the parameter updating of the other factor. However, including those effects in the present application do not improve the model performance.

described above. I therefore fit a regression

$$\text{DAPM}: \qquad \lambda_T^j = \lambda_0 + \Lambda_1^{j,fd}|\overline{fd}_t| + \Lambda_1^{j,\Delta VOL}\Delta VOL_t + \Lambda_1^{j,\Delta CRB}\Delta CRB_t + \Lambda_1^{j,\Delta Cr}\Delta Cr_t + \Lambda_1^{j,\Delta FRB}\Delta FRB_t \tag{39}$$

with the method of Adrian et al. (2015) for each of the two cross-sectional risk factors $j = DOL, HML$.

[Table 3 about here.]

Table 3 provides estimates of risk price parameters from the SLAPM model as well as the two benchmark specifications. The first two columns show the parameter estimates of the SLAPM model along with standard errors in parenthesis. Latter are calculated by inverting a numerically computed Fisher information matrix as described in section 2.3. The column shows positive estimates of the parameters a_{DOL} and a_{HML} that are significant on a one percent level. Hence, the contemporaneous model errors s_t significantly impact their corresponding price of risk and are therefore informative for inferring variation in risk premia. The autoregressive parameters b_{DOL} and b_{HML} are both significant but fairly moderate in magnitude. Both factor risk price processes are therefore crucially varying with a moderate level of persistence. The third and fourth column show the estimation result for the DAPM. Shocks to FX volatility have a significant impact on both risk prices (at least on a five percent significance level), whereas the other factors either forecast DOL or HML risk prices. The fifth and sixth column show constant risk price estimates when assuming no time-variation. In line with Lustig et al. (2011), the DOL risk price is found close to zero, whereas the HML risk price is significant and close to unity.

[Figure 3 about here.]

I turn next to the currency risk premia series filtered from the two dynamic approaches under investigation. These premia are computed via the conditional expectation of the corresponding currency trading strategy. More precisely, the DOL and HML risk premium series are derived as

$$\hat{rp}_{t+1}^{DOL} := \hat{\mathbb{E}}_t\left(DOL_{t+1}\right) = \hat{\mathbb{E}}_t\left(\frac{1}{5}\sum_{i=1}^{5} r_{t+1}^{Ci}\right) \tag{40}$$

$$= \frac{1}{5}\sum_{i=1}^{5}\left(\hat{\beta}_{Ci}^{DOL}\hat{\lambda}_t^{DOL} + \hat{\beta}_{Ci}^{HML}\hat{\lambda}_t^{HML}\right) \tag{41}$$

and

$$\hat{rp}_t^{HML} := \hat{\mathbb{E}}_t\left(HML_{t+1}\right) = \hat{\mathbb{E}}_t\left(r_{t+1}^{C5} - r_{t+1}^{C1}\right) \tag{42}$$

$$= (\hat{\beta}_{C5}^{DOL} - \hat{\beta}_{C1}^{DOL})\hat{\lambda}_t^{DOL} + (\hat{\beta}_{C5}^{HML} - \hat{\beta}_{C1}^{HML})\hat{\lambda}_t^{HML} \tag{43}$$

where $\hat{\mathbb{E}}_t$ represents the conditional mean operator implied by the corresponding method.

Figure 3 shows filtered risk premia implied by the SLAPM (continuous line) and the DAPM (dotted line). Consider the filtered DOL risk premia series in panel (a) first. Both approaches find a series that is fluctuating around some level just above zero showing negative as well as positive values. This variation may explain that the DOL risk premium, although providing reasonable explanatory power in the cross-section of carry trade returns (Lustig et al. (2011), is usually estimated with a insignificant price of risk. The DAPM-filtered risk premia show more pronounced swings in general and a substantial downward spike during the global financial crisis that is much less pronounced in the SLAPM-filtered series.

The filtered carry premia shown in panel (b) show more persistent dynamics with series implied by the two competing dynamic approaches run quite close to each other. Both series capture periods of relatively high premia in the 2000s and after 2015 in line with the increased profitability of carry trades in these times. However, there are some differences. First, the SLAPM-implied carry trade premia show more substantial drops in which the expected carry return even turns negative though with fast recovery. Most of these crucial drops do not refer to recession periods (with the recent global financial crisis being an exception) but fall together with huge currency movements disconnected from macroeconomic shocks, as have been already documented for carry trade returns by Brunnermeier et al. (2008). One example of such a disconnected currency movement is the strong appreciation of the Japanese yen, a typical short-leg carry currency, in October 1998, accompanied by a huge drop in the SLAPM-implied carry return premium. In line with the observed disconnect, the downward spikes are not entirely captured by the DAPM-implied premia that infer risk price movements from economic forecast variables. An exception is the drop in the global financial crises wich is estimated with almost the same magnitude by both approaches. A second difference is the interpretation of the recent carry return race that started around 2015 as can be seen in Figure 2. Whereas the DAPM refer to these returns as being expected due to the higher estimated risk premia, the SLAPM propagates less increased risk premia and devotes a higher fraction of realized carry return to idiosyncratic innovations.

The next step is to achieve insights whether the SLAPM shows a better picture of the actual risk price dynamics or it is just unable to retrieve the same amount of information about risk price dynamics that the DAPM is capable of due to its time series model.

[Figure 4 about here.]

The unconditional model fits the carry trade cross-section quite well already. This is indicated by Figure 4(a) that shows a scatter plot of realized average returns against average returns from the fitted unconditional model. Hence, their is not much potential for improvement in the model performance in fitting average portfolio returns. However, whereas the DAPM, the scatter plot of which is shown in Figure 4(b), can not improve the fit over the unconditional model, the SLAPM appears to improve the average fit of the low interest rate portfolio C1, as indicated in Figure 4(c).

Dynamic models are unlikely to show their merits by comparing time averages of returns. I therefore turn to a comparison of pricing errors. The two metrics being considered are the root mean squared error (RMSE) computed for each test asset i as

$$RMSE_i = \sqrt{\frac{1}{T-1}\sum_{t=2}^{T} \hat{e}_{i,t}^2} \tag{44}$$

with $\hat{e}_t^2$ being the fitted residuals from the corresponding model and the root mean squared forecast error (RMSFE), computed as

$$RMSFE_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T-1} \left(r_{t+1} - \hat{\beta}_i^{DOL}\hat{\lambda}_t^{DOL} - \hat{\beta}_i^{HML}\hat{\lambda}_t^{HML}\right)^2}. \tag{45}$$

In addition to the performance of the SLAPM and the two benchmarks, I evaluate the pricing ability of the SLAPM with risk prices shifted one period backward and refer it as "in-sample" SLAPM. Returns for period $t+1$ are usually predicted using the time t estimate of the risk price λ_t and the corresponding exposures. Remark that λ_{t+1} is by construction updated towards the optimal risk price given the return and cross-sectional factor realizations in time $t+1$. Hence, an improved ex-post prediction for r_{t+1} would be based on λ_{t+1} instead of λ_t. This ex-post prediction is not available in a out-of-sample exercise but would be the natural candidate when filtering risk prices in-sample.

[Table 4 about here.]

Table 4 shows the RMSE and RMSFE of the both SLAPM variants and the two benchmarks. Let us concentrate on the improvement of incorporating economic predictors first. In the first two rows of panel (a) we observe that the RMSE of the DAPM compared to the unconditional model is lower for every portfolio, except C2. The improved pricing ability is particularly evident for the extreme portfolios and returns to the two factors DOL and HML. The third row in panel (a) shows the RMSEs from the SLAPM model. The RMSE regarding the HML factor returns is smaller than that for the unconditional model but slightly higher compared to the DAPM. Striking is that the SLAPM generates even lower RMSE's for the low interest portfolio C1, as suspected from visual inspection with the scatter plots, and the

DOL factor returns. However, the SLAPM is less successful in pricing the cross-section with a RMSE of 0.880 on average. Both, DAPM and the in-sample SLAPM achieve a RMSE of about 0.849. A comparison of RMSFEs is shown in panel (b) of table (6). Here, the SLAPM RMSFEs, except for C2, lie between those of the unconditional model and the DAPM. The SLAPM forecasts are constructed such that the one-period ahead prediction error is minimized at the current observation. This construction yields the RMSFEs of the in-sample SLAPM predictions are well below those of the DAPM and the unconditional model.

In sum, Table 4 documents that the SLAPM, inferring information from the cross-sectional model only, can indeed compete with the DAPM employing economic time series predictors with regard to mean squared (forecast) errors. The results suggest that the DAPM is more capable in forecasting. This is because it exploits information in economic forecast variables that becomes available to the SLAPM with a delay. From an in-sample perspective, the SLAPM appears to track movements in risk prices more adequate. This particularly holds for the DOL risk price that is estimated with substantially lower mean squared errors. Hence, the DOL risk premium for investing in the whole currency market against the U.S. dollar appears to fluctuate less than predicted with economic forecast variables. Results concerning HML premium for investing in a long-short carry trade suggest that the economically motivated forecast model captures most of the predictable components that are exploited for global currency hedging in Opie and Riddiough (2020). However, these forecast variables do not entirely capture risk price movements due to currency crashes disconnected from macroeconomic conditions that are suspected being the cause for carry trade premia as argued by Brunnermeier et al. (2008). In line with the idea that carry premia are a copensation for crash risk, carry premia depress after crash occurrence and build up in the aftermath until the next crash occurs. This is reasonable if expecting crashes being unlikely to occur right after each other. Of course, the SLAPM can not predict these crashes either, but takes them into account in in-sample filtering of risk premia.

5 Conclusions

The paper introduced an empirical dynamic asset pricing framework that allows for time-varying lambdas and betas which are unobserved processes filtered from the cross-section of asset returns and the asset pricing model's factor structure in line with the more general GAS model from Creal et al. (2013). It is applicable to a wide range of linear factor models in the finance literature. A main advantage is that no forecasters or instruments are required to describe the time dynamics of risk prices or exposures. A Monte-Carlo study provides evidence that the method is capable of filtering substantial risk price movements from a model with correctly specified cross-sectional factors under a realistic signal to noise ratio. Moreover, it can compete with a dynamic estimation approach taking signals of true time series drivers into account. The results point towards a non-negligible source of possible misspecification in the time series model that can be circumvented by employing the SLAPM framework.

Updating schemes for the SLAPM class with elliptically distributed innovations have been derived. It turns out that the risk price updating direction within this class is unaffected by the distributional assumptions but the magnitude of the price movement depends on the shape of the corresponding probability density function. A particular tractable model with respect to complexity and computational burden for estimation is the presented SLAPM(1,1) specification with constant betas and normally distributed innovation terms. Its use have been illustrated by an empirical application filtering adequate series of currency risk premia that reveal movements not displayed by methods using economic forecast factors.

References

Adrian, T., Crump, R. K., Moench, E., 2015. Regression-based Estimation of Dynamic Asset Pricing Models. Journal of Financial Economics 118 (2), 211–244.

Adrian, T., Crump, R. K., Vogt, E., 2019. Nonlinearity and Flight-to-Safety in the Risk-Return Trade-Off for Stocks and Bonds. Journal of Finance 74 (4), 1931–1973.

Adrian, T., Etula, E., Groen, J. J. J., 2011. Financial Amplification of Foreign Exchange Risk Premia. European Economic Review 55 (3), 354–370.

Avdjiev, S., Du, W., Koch, C., Shin, H. S., 2019. The dollar, bank leverage, and deviations from covered interest parity. American Economic Review: Insights 1 (2), 193–208.

Bakshi, G., Panayotov, G., 2013. Predictability of Currency Carry Trades and Asset Pricing Implications. Journal of Financial Economics 110 (1), 139–163.

Bansal, R., Yaron, A., 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. Journal of Finance 59 (4), 1481–1509.

Bernardi, M., Catania, L., 2019. Switching Generalized Autoregressive Score Copula Models with Application to Systemic Risk. Journal of Applied Econometrics 34 (1), 43–65.

Bhansali, V., 2007. Volatility and the Carry Trade. The Journal of Fixed Income 17, 72–84.

Blasques, F., Koopman, S. J., Lucas, A., 2014. Maximum Likelihood Estimation for Score-Driven Models. Tinbergen Institute Discussion Paper, TI 14-029/III.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31 (3), 307–327.

Braun, P. A., Nelson, D. B., Sunier, A. M., 1995. Good News, Bad News, Volatility, and Betas. Journal of Finance 50 (5), 1575–1603.

Brunnermeier, M. K., Nagel, S., Pedersen, L. H., 2008. Carry Trades and Currency Crashes. NBER Macroeconomics Annual 23 (1), 313–348.

Brunnermeier, M. K., Pedersen, L. H., 2009. Market Liquidity and Funding Liquidity. Review of Financial studies 22 (6), 2201–2238.

Campbell, J. Y., 1987. Stock Returns and the Term Structure. Journal of Financial Economics 18 (2), 373–399.

Campbell, J. Y., Shiller, R. J., 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. Review of Financial Studies 1 (3), 195–228.

Campbell, J. Y., Thompson, S. B., 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21 (4), 1509–1531.

Campbell, J. Y., Yogo, M., 2006. Efficient Tests of Stock Return Predictability. Journal of Financial Economics 81 (1), 27–60.

Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. Journal of Finance 52 (1), 57–82.

Chaieb, I., Langlois, H., Scaillet, O., 2018. Time-Varying Risk Premia in Large International Equity Markets. Swiss Finance Institute Research Paper No. 18-04.

Chen, Y.-c., Rogoff, K., 2003. Commodity Currencies. Journal of International Economics 60 (1), 133–160.

Cochrane, J. H., 2008. The Dog That Did Not Bark: A Defense of Return Predictability. Review of Financial Studies 21 (4), 1533–1575.

Cochrane, J. H., 2011. Discount Rates. Journal of Finance 66 (4), 1047–1108.

Creal, D., Koopman, S. J., Lucas, A., 2013. Generalized Autoregressive Score Models with Applications. Journal of Applied Econometrics 28 (5), 777–795.

Diebold, F. X., Mariano, R. S., 1995. Comparing Predictive Accuracy. Journal of Business & Economic Statistics 13 (3), 253–263.

Embrechts, P., Frey, R., McNeil, A. J., 2015. Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition. Princeton University Press.

Engle, R. F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50 (4), 987–1007.

Fama, E. F., French, K. R., 1989. Business Conditions and Expected Returns on Stocks and Bonds. Journal of Financial Economics 25 (1), 23–49.

Fama, E. F., French, K. R., 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics 33 (1), 3–56.

Fama, E. F., MacBeth, J. D., 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 81 (3), 607–636.

Fang, K.-T., Kotz, S., Ng, K. W., 1990. Symmetric Multivariate and Related Distributions. Chapman & Hall/CRC, London ; New York.

Ferson, W. E., Harvey, C. R., 1991. The Variation of Economic Risk Premiums. Journal of Political Economy 99 (2), 385–415.

Gagliardini, P., Ossola, E., Scaillet, O., 2016. Time-Varying Risk Premium in Large Cross-Sectional Equity Data Sets. Econometrica 84 (3), 985–1046.

Gagliardini, P., Ossola, E., Scaillet, O., 2019. Estimation of Large Dimensional Conditional Factor Models in Finance. Swiss Finance Institute Research Paper No. 19-46.

Ghysels, E., 1998. On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt? Journal of Finance 53 (2), 549–573.

Gorgi, P., Hansen, P. R., Janus, P., Koopman, S. J., 2019. Realized Wishart-GARCH: A Score-driven Multi-Asset Volatility Model. Journal of Financial Econometrics 17 (1), 1–32.

Harvey, A., Lange, R.-J., 2017. Volatility Modeling with a Generalized t Distribution. Journal of Time Series Analysis 38 (2), 175–190.

Harvey, A. C., May 2013. Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series, new Edition. Cambridge University Press, Cambridge ; New York.

He, Z., Kelly, B., Manela, A., 2017. Intermediary Asset Pricing: New Evidence from many Asset Classes. Journal of Financial Economics 126 (1), 1–35.

Jagannathan, R., Wang, Z., 1996. The Conditional CAPM and the Cross-Section of Expected Returns. Journal of Finance 51 (1), 3–53.

Keim, D., Stambaugh, R., 1986. Predicting Returns in the Stock and Bond Markets. Journal of Financial Economics 17 (2), 357–390.

Lettau, M., Ludvigson, S., 2001. Consumption, Aggregate Wealth, and Expected Stock Returns. Journal of Finance 56 (3), 815–849.

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

Lustig, H., Roussanov, N., Verdelhan, A., 2014. Countercyclical currency risk premia. Journal of Financial Economics 111 (3), 527–553.

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

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

Oh, D. H., Patton, A. J., 2018. Time-Varying Systemic Risk: Evidence From a Dynamic Copula Model of CDS Spreads. Journal of Business & Economic Statistics 36 (2), 181–195.

Opie, W., Riddiough, S. J., 2020. Global Currency Hedging with Common Risk factors. Journal of Financial Economics 136 (3), 780–805.

Rossi, B., 2013. Exchange Rate Predictability. Journal of Economic Literature 51 (4), 1063–1119.

Appendix

The following technical derivations prove the main results of SLAPM models with elliptically distributed innovations. To be clear about the setting in which the following results hold the elliptical SLAPM(p,q)-model is summarized in the following definition:

Definition 1. Let $(r_t)_{t=1}^T$ and $(f_t)_{t=1}^T$ time series of dimensions N and K respectively. We define the **elliptical SLAPM(p,q)-model** as the system given by

$$f_{t+1} = \mu + \Phi f_t + u_{t+1} \tag{46}$$

$$r_{t+1} = \beta_t \lambda_t + \beta_t u_{t+1} + e_{t+1} \tag{47}$$

$$\theta_t = \omega + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_j \theta_{t-j}, \qquad \theta_t = \begin{pmatrix} \lambda_t \\ vec(\beta_t) \end{pmatrix} \tag{48}$$

$$s_t = \begin{pmatrix} s_t^\lambda \\ s_t^\beta \end{pmatrix} = \begin{pmatrix} \mathcal{I}_t^\lambda & \mathcal{I}_t^{\lambda,\beta} \\ \mathcal{I}_t^{\beta,\lambda} & \mathcal{I}_t^\beta \end{pmatrix}^{-1} \begin{pmatrix} \nabla_t^\lambda \\ \nabla_t^\beta \end{pmatrix} = \mathcal{I}_t^{-1} \nabla_t \tag{49}$$

$$p_u(u_t) = |\Omega_u|^{-\frac{1}{2}} \psi^u \left(u_t' \Omega_u^{-1} u_t, \nu_u \right), \qquad p_e(e_t) = |\Omega_e|^{-\frac{1}{2}} \psi^e \left(e_t' \Omega_e^{-1} e_t, \nu_e \right) \tag{50}$$

with observational density score ∇_t, associated Fisher information matrix $\mathcal{I}_t^{-1}$, elliptical density generators $\psi^e, \psi^u : \mathbb{R}_{\geq 0} \times \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ and parameter vectors/matrices $\mu, \Phi, \omega, A_1, \ldots, A_p, B_1, \ldots, B_q, \Omega_u$, Ω_e, ν_u, ν_e of appropriate dimensions.

Before dealing with the main technical results of the paper, I provide a lemma collecting some helpful properties of spherically distributed random vectors. A n-dimensional random vector x is spherically distributed if $x = rs$, where s is uniformly distributed on the $(n-1)$-dimensional unit sphere and r is a non-negative random number that is independent of s (Fang et al. (1990)). Moreover, let $\|x\|$ denote the euclidean norm of an random vector x.

Lemma 1. *Suppose the n-dimensional random vector x follows a spherical distribution, then:*

(i) x *and* $\frac{x}{\|x\|}$ *are independent.*

(ii) $E\left(\frac{x}{\|x\|} \frac{x^\top}{\|x\|} \right) = \frac{1}{n} I_n$

Proof. See proofs of Theorem 2.3 and Theorem 2.7 in Fang et al. (1990). □

We now turn to the derivation of general formulas for the scores and information matrices for SLAPM models with elliptical distribution assumptions represented by density generator functions ψ^e and ψ^u. An explicit parameter updating scheme in the case of zero cross-information quantities ($\mathcal{I}_t^{\lambda,\beta} = \mathcal{I}_t^{\beta,\lambda} = 0$) will be derived in Corollary 2. Further results provide scores and information quantities for Gaussian and Student's-t SLAPM models.

Proposition 1. The score and corresponding Fisher information matrix in the updating scheme of an elliptical SLPAM(p,q)-model are given by

$$\nabla_t = \begin{pmatrix} \nabla_t^\lambda \\ \nabla_t^\beta \end{pmatrix} = -2 \frac{\psi_1^e}{\psi^e} \begin{pmatrix} \beta_{t-1}^\top \Omega_e^{-1} e_t \\ (\lambda_{t-1} + u_t) \otimes \Omega_e^{-1} e_t \end{pmatrix} \tag{51}$$

and

$$
\begin{aligned}
\mathcal{I}_t &= \begin{pmatrix} \mathcal{I}_t^{\lambda} & \mathcal{I}_t^{\lambda,\beta} \\ \mathcal{I}_t^{\beta,\lambda} & \mathcal{I}_t^{\beta} \end{pmatrix} \\
&= C_{\mathcal{I}}(\psi^e) \begin{pmatrix} \beta_{t-1}^{\top}\Omega_e^{-1}\beta_{t-1} & \lambda_{t-1}^{\top} \otimes \beta_{t-1}^{\top}\Omega_e^{-1} \\ \lambda_{t-1} \otimes \Omega_e^{-1}\beta_{t-1} & ((\lambda_{t-1}\lambda_{t-1}^{\top} + \mathbb{E}\left(u_t u_t^{\top}\right)) \otimes \Omega_e^{-1} \end{pmatrix}
\end{aligned} \tag{52}
$$

with

$$
C_{\mathcal{I}}(\psi^e) = \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2 \left(\frac{\psi_1^e}{\psi^e}\right)^2 \right) \tag{53}
$$

where $\tilde{e}_t = \Omega_e^{-1/2} e_t$ and ψ_1^e is the partial derivative with respect to the first component.

Proof. The common observational density of u_t and e_t can be derived with the given density functions. We find

$$
\begin{aligned}
p(e_t, u_t|\mathcal{F}_{t-1}, \theta_t) &= p(e_t|u_t, \mathcal{F}_{t-1}, \theta_t)p(u_t|\mathcal{F}_{t-1}, \theta_t) \\
&= |\Omega_e|^{-\frac{1}{2}}\psi^e\left(e_t^{\top}\Omega_e^{-1}e_t, \nu_e\right)|\Omega_u|^{-\frac{1}{2}}\psi^u\left(u_t^{\top}\Omega_u^{-1}u_t, \nu_u\right)
\end{aligned}
$$

and the log-likelihood

$$
\begin{aligned}
l_t &= \ln p(e_t, u_t|\mathcal{F}_{t-1}, \theta_t) \\
&= -\frac{1}{2}\ln|\Omega_e| - \frac{1}{2}\ln|\Omega_u| + \ln\psi^e\left(e_t^{\top}\Omega_e^{-1}e_t, \nu_e\right) + \ln\psi^u\left(u_t^{\top}\Omega_u^{-1}u_t, \nu_u\right).
\end{aligned}
$$

The components of the score can be computed as

$$
\begin{aligned}
\frac{\partial l_t}{\partial\theta} &= \frac{\partial}{\partial\theta}\ln\psi^e\left(e_t^{\top}\Omega_e^{-1}e_t, \nu_e\right) \\
&= \frac{\psi_1^e}{\psi^e}\frac{\partial}{\partial\theta}\left(e_t^{\top}\Omega_e^{-1}e_t\right) \\
&= 2\frac{\psi_1^e}{\psi^e}\frac{\partial e_t}{\partial\theta}\Omega_e^{-1}e_t
\end{aligned}
$$

with ψ_1^e being the first derivative of ψ with respect to its first component. The score functions of risk prices and exposures can then be obtained as follows:

$$
\nabla_t^{\lambda} = \frac{\partial l_t}{\partial\lambda_{t-1}} = 2\frac{\psi_1^e}{\psi^e}\frac{\partial e_t}{\partial\lambda_{t-1}}\Omega_e^{-1}e_t = -2\frac{\psi_1^e}{\psi^e}\beta_{t-1}^{\top}\Omega_e^{-1}e_t
$$

and

$$
\begin{aligned}
\nabla_t^{\beta} &= \frac{\partial l_t}{\partial vec(\beta_{t-1})} = 2\frac{\psi_1^e}{\psi^e}\frac{\partial e_t}{\partial vec(\beta_{t-1})}\Omega_e^{-1}e_t \\
&= 2\frac{\psi_1^e}{\psi^e}\left(\frac{\partial}{\partial vec(\beta_{t-1})}(r_t - \beta_{t-1}(\lambda_{t-1}+u_t))\right)\Omega_e^{-1}e_t \\
&= -2\frac{\psi_1^e}{\psi^e}\left(\frac{\partial}{\partial vec(\beta_{t-1})}(vec(\beta_{t-1}(\lambda_{t-1}+u_t)))\right)\Omega_e^{-1}e_t \\
&= -2\frac{\psi_1^e}{\psi^e}\left(\frac{\partial}{\partial vec(\beta_{t-1})}(((\lambda_{t-1}+u_t)^\top \otimes I_N)vec(\beta_{t-1}))\right)\Omega_e^{-1}e_t \\
&= -2\frac{\psi_1^e}{\psi^e}((\lambda_{t-1}+u_t)\otimes I_N)\Omega_e^{-1}e_t \\
&= -2\frac{\psi_1^e}{\psi^e}((\lambda_{t-1}+u_t)\otimes I_N)(1\otimes\Omega_e^{-1}e_t) \\
&= -2\frac{\psi_1^e}{\psi^e}((\lambda_{t-1}+u_t)\otimes\Omega_e^{-1}e_t).
\end{aligned}
$$

Define the normalized error $\tilde{e}_t = \Omega_e^{-1/2}e_t$ with $\Omega_e^{1/2}$ being the Cholesky factor of the corresponding dispersion matrix. Remark that $\tilde{e}_t$ is spherically distributed and therefore fulfills the conditions for applying Lemma 1. The Fisher information matrix can then be computed as

$$
\begin{aligned}
\mathcal{I}_t &= \mathbb{E}_{t-1}\left(\frac{\partial l_t}{\partial\theta}\frac{\partial l_t^\top}{\partial\theta}\right) \\
&= \mathbb{E}_{t-1}\left(\left(2\frac{\psi_1^e}{\psi^e}\right)^2\frac{\partial e_t}{\partial\theta}\Omega_e^{-1}e_t(\Omega_e^{-1}e_t)^\top\frac{\partial e_t^\top}{\partial\theta}\right) \\
&= \mathbb{E}_{t-1}\left(\left(2\frac{\psi_1^e}{\psi^e}\right)^2\frac{\partial e_t}{\partial\theta}\Omega_e^{-\frac{1}{2}}\tilde{e}_t\tilde{e}_t^\top\Omega_e^{-\frac{1}{2}}\frac{\partial e_t^\top}{\partial\theta}\right) \\
&= 4\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1}{\psi}\right)^2\frac{\partial e_t}{\partial\theta}\Omega_e^{-\frac{1}{2}}\mathbb{E}_{t-1}\left(\frac{\tilde{e}_t}{\|\tilde{e}_t\|}\left(\frac{\tilde{e}_t}{\|\tilde{e}_t\|}\right)^\top \,\middle|\, u_t\right)\Omega_e^{-\frac{1}{2}}\frac{\partial e_t^\top}{\partial\theta}\right) \\
&= \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1}{\psi}\right)^2\frac{\partial e_t}{\partial\theta}\Omega_e^{-\frac{1}{2}}\Omega_e^{-\frac{1}{2}}\frac{\partial e_t^\top}{\partial\theta}\right) \\
&= \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1}{\psi}\right)^2\right)\mathbb{E}_{t-1}\left(\frac{\partial e_t}{\partial\theta}\Omega_e^{-1}\frac{\partial e_t^\top}{\partial\theta}\right) \\
&= C_{\mathcal{I}}(\psi^e)\ \mathbb{E}_{t-1}\left(\frac{\partial e_t}{\partial\theta}\Omega_e^{-1}\frac{\partial e_t^\top}{\partial\theta}\right)
\end{aligned}
$$

with $C_{\mathcal{I}}(\psi^e) = \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1^e}{\psi^e}\right)^2\right)$. The fourth and fifth equality holds because of Lemma 1(i) and (ii), respectively.[17] Information quantities with respect to factor risk prices and exposures can

[17] $\frac{\partial e_t}{\partial\theta}$ being independent of e_t is exploited as well.

then be computed as follows:

$$\mathcal{I}_t^{\lambda} = C_{\mathcal{I}}(\psi^e)\beta_{t-1}^{\top}\Omega_e^{-1}\beta_{t-1}$$

$$\begin{aligned}
\mathcal{I}_t^{\beta} &= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left(((\lambda_{t-1}+u_t)\otimes I_N)\Omega_e^{-1}((\lambda_{t-1}+u_t)^{\top}\otimes I_N)\right)\\
&= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left(((\lambda_{t-1}+u_t)\otimes I_N)(1\otimes\Omega_e^{-1})((\lambda_{t-1}+u_t)^{\top}\otimes I_N)\right)\\
&= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left(((\lambda_{t-1}+u_t)\otimes\Omega_e^{-1})((\lambda_{t-1}+u_t)^{\top}\otimes I_N)\right)\\
&= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left(((\lambda_{t-1}+u_t)(\lambda_{t-1}+u_t)^{\top})\otimes\Omega_e^{-1}\right)\\
&= C_{\mathcal{I}}(\psi^e)(\lambda_{t-1}\lambda_{t-1}^{\top}+\mathbb{E}(u_tu_t^{\top}))\otimes\Omega_e^{-1}
\end{aligned}$$

$$\begin{aligned}
\left(\mathcal{I}_t^{\lambda,\beta}\right)^{\top} = \mathcal{I}_t^{\beta,\lambda} &= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left(((\lambda_{t-1}+u_t)\otimes I_N)\Omega_e^{-1}\beta_{t-1}\right)\\
&= C_{\mathcal{I}}(\psi^e)\mathbb{E}_{t-1}\left((\lambda_{t-1}+u_t)\otimes\Omega_e^{-1}\beta_{t-1}\right)\\
&= C_{\mathcal{I}}(\psi^e)(\lambda_{t-1}\otimes\Omega_e^{-1}\beta_{t-1}).
\end{aligned}$$

□

Corollary 2. *The scaled scores of an SLAPM(p,q) with characteristic distribution generator* ψ^e *and assuming* $(\mathcal{I}_t^{\lambda,\beta})^{\top} = \mathcal{I}_t^{\beta,\lambda} \equiv 0$ *is given by*

$$\begin{pmatrix} s_t^{\lambda} \\ s_t^{\beta} \end{pmatrix} = C(\tilde{e}_t,\psi^e)\begin{pmatrix} (\beta_{t-1}^{\top}\Omega_e^{-1}\beta_{t-1})^{-1}\beta_{t-1}^{\top}\Omega_e^{-1}r_t-(\lambda_{t-1}+u_t) \\ vec(e_t(\lambda_{t-1}+u_t)^{\top}(\lambda_{t-1}\lambda_{t-1}^{\top}+\mathbb{E}_{t-1}(u_tu_t^{\top}))) \end{pmatrix}$$

with

$$C(\tilde{e}_t,\psi^e) = -\frac{2\psi_1^e}{\psi^e}C_{\mathcal{I}}(\psi^e)^{-1} = -\frac{N\frac{\psi_1^e}{\psi^e}}{2\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1^e}{\psi^e}\right)^2\right)}.$$

Proof. Employing the results from Proposition 1 while assuming zero cross information quantities i.e. $\mathcal{I}_t^{\lambda,\beta} = \mathcal{I}_t^{\beta,\lambda} = 0$ we can derive the following formulas for the driving martingale difference sequences of time-varying risk prices and exposures. For λ_t we find:

$$\begin{aligned}
s_t^{\lambda} &= (\mathcal{I}_t^{\lambda})^{-1}\nabla_t^{\lambda}\\
&= -2\frac{\psi_1^e}{\psi^e}C_{\mathcal{I}}(\psi^e)^{-1}(\beta^{\top}\Omega_e^{-1}\beta)^{-1}\beta^{\top}\Omega_e^{-1}(e_t))\\
&= C(\tilde{e}_t,\psi^e)(\beta^{\top}\Omega_e^{-1}\beta)^{-1}\beta^{\top}\Omega_e^{-1}(r_t-\beta(\lambda_{t-1}+u_t))\\
&= C(\tilde{e}_t,\psi^e)\left[(\beta^{\top}\Omega_e^{-1}\beta)^{-1}\beta^{\top}\Omega_e^{-1}r_t-(\lambda_{t-1}+u_t)\right]. \qquad (54)
\end{aligned}$$

For β_t, we analogously find:

$$
\begin{aligned}
s_t^{\beta} &= (\mathcal{I}_t^{\beta})^{-1}\nabla_t^{\beta} \\
&= -2\frac{\psi_1^e}{\psi^e} C_{\mathcal{I}}(\psi^e)^{-1}((\lambda_{t-1}\lambda_{t-1}^\top + \mathbb{E}_{t-1}(u_t u_t^\top)) \otimes \Omega_e^{-1})^{-1}((\lambda_{t-1} + u_t) \otimes \Omega_e^{-1} e_t) \\
&= C(\tilde{e}_t, \psi^e)((\lambda_{t-1}\lambda_{t-1}^\top + \mathbb{E}_{t-1}(u_t u_t^\top)) \otimes \Omega_e^{-1})^{-1}((\lambda_{t-1} + u_t) \otimes \Omega_e^{-1}) e_t \\
&= C(\tilde{e}_t, \psi^e)(((\lambda_{t-1}\lambda_{t-1}^\top + \mathbb{E}_{t-1}(u_t u_t^\top))(\lambda_{t-1} + u_t)) \otimes I_N) vec(e_t) \\
&= C(\tilde{e}_t, \psi^e) vec(e_t(\lambda_{t-1} + u_t)^\top (\lambda_{t-1}\lambda_{t-1}^\top + \mathbb{E}_{t-1}(u_t u_t^\top)))
\end{aligned}
$$

□

Corollary 3 (Gaussian Residuals)**.** *The score and the corresponding Fisher information matrix of the elliptical SLAPM(p,q)-model with Gaussian residuals are given by*

$$
\nabla_t = \begin{pmatrix} \beta_{t-1}^\top \Omega_e^{-1} e_t \\ (\lambda_{t-1} + u_t) \otimes \Omega_e^{-1} e_t \end{pmatrix}
$$

and

$$
\mathcal{I}_t = \begin{pmatrix} \beta_{t-1}^\top \Omega_e^{-1} \beta_{t-1} & \lambda_{t-1}^\top \otimes \beta_{t-1}^\top \Omega_e^{-1} \\ \lambda_{t-1} \otimes \Omega_e^{-1} \beta_{t-1} & ((\lambda_{t-1}\lambda_{t-1}^\top + \Omega_u) \otimes \Omega_e^{-1} \end{pmatrix}.
$$

Proof. The characteristic generator of the multivariate normal density is given by $\psi^e(x, \nu) = (2\pi)^{N/2} e^{-x/2}$. By observing that ψ^e solves the differential equation $\psi_1^e = -\frac{1}{2}\psi^e$, we find that $-2\psi_1^e/\psi^e = 1$ and

$$
\begin{aligned}
C_{\mathcal{I}}(\tilde{e}_t, \psi^e) &= \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2 \left(\frac{\psi_1^e}{\psi^e}\right)^2\right) = \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2 \left(\frac{-\psi^e/2}{\psi^e}\right)^2\right) \\
&= \frac{1}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\right) = \frac{1}{N}\mathbb{E}_{t-1}\left(\tilde{e}_t^\top \tilde{e}_t\right) = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}_{t-1}\left(\tilde{e}_{it}^\top \tilde{e}_{it}\right) = 1
\end{aligned}
$$

because $\tilde{e}_t \sim N(0, I_N)$. Moreover, normally distributed factor innovations yield $\mathbb{E}\left(u_t u_t^\top\right) = \Omega_u$. The assertion now follows with Proposition 1.

□

Corollary 4 (Student's t Residuals)**.** *The score and the corresponding Fisher information matrix of the elliptical SLAPM(p,q) with residuals following a Student's t distribution are given by*

$$
\nabla_t = \frac{\nu_e + N}{\nu_e + e_t^\top \Omega_e^{-1} e_t} \begin{pmatrix} \beta_{t-1}^\top \Omega_e^{-1} e_t \\ (\lambda_{t-1} + u_t) \otimes \Omega_e^{-1} e_t \end{pmatrix}
$$

and

$$
\mathcal{I}_t = \frac{\nu_e + N}{\nu_e + N + 2} \begin{pmatrix} \beta_{t-1}^\top \Omega_e^{-1} \beta_{t-1} & \lambda_{t-1}^\top \otimes \beta_{t-1}^\top \Omega_e^{-1} \\ \lambda_{t-1} \otimes \Omega_e^{-1} \beta_{t-1} & ((\lambda_{t-1}\lambda_{t-1}^\top + \frac{\nu_u}{\nu_u - 2}\Omega_u) \otimes \Omega_e^{-1} \end{pmatrix}.
$$

Proof. The characteristic generator of the multivariate Student's t density is given by

$$\psi(x,\nu_e) = \frac{\Gamma\left(\frac{\nu_e+N}{2}\right)}{(\nu_e\pi)^{\frac{N}{2}}\Gamma\left(\frac{\nu_e}{2}\right)}\left(1+\frac{x}{\nu_e}\right)^{-\frac{\nu_e+N}{2}} \tag{55}$$

and can be used to compute

$$-2\frac{\psi_1^e}{\psi^e}(e_t^\top\Omega_e^{-1}e_t,\nu_e) = -2\frac{-\frac{\nu_e+N}{2}\left(1+\frac{e_t^\top\Omega_e^{-1}e_t}{\nu_e}\right)^{-\frac{\nu_e+N}{2}-1}\frac{1}{\nu_e}}{\left(1+\frac{e_t^\top\Omega_e^{-1}e_t}{\nu_e}\right)^{-\frac{\nu_e+N}{2}}} = \frac{\nu_e+N}{\nu_e+e_t^\top\Omega_e^{-1}e_t}. \tag{56}$$

The scaling term of the information quantity can be reformulated as follows:

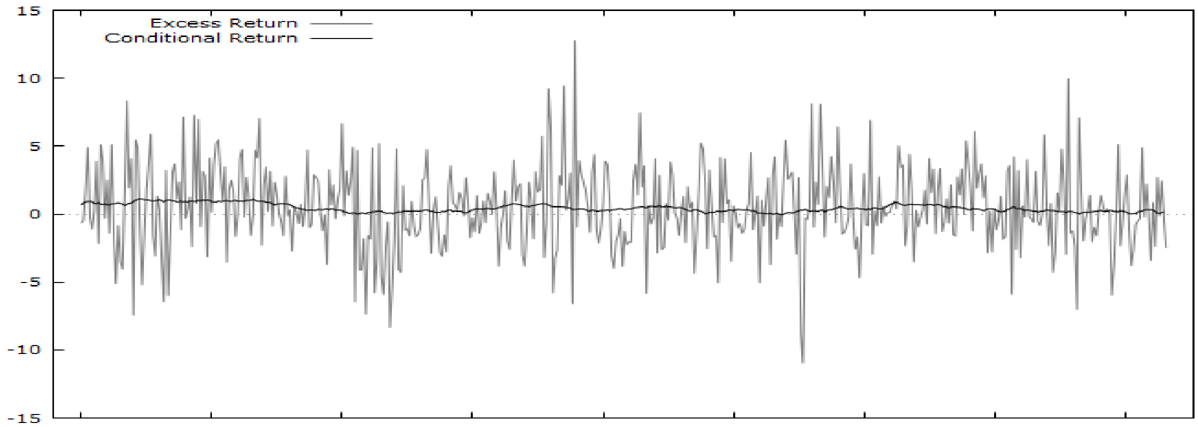
$$\begin{aligned} C_{\mathcal{I}}(\tilde{e}_t,\psi^e) &= \frac{4}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\psi_1^e}{\psi^e}\right)^2\right) = \frac{1}{N}\mathbb{E}_{t-1}\left(\|\tilde{e}_t\|^2\left(\frac{\nu_e+N}{\nu_e+\|\tilde{e}_t\|^2}\right)^2\right) \\ &= \left(\frac{\nu_e+N}{N}\right)^2\mathbb{E}_{t-1}\left(\frac{Z}{\left(\frac{\nu_e}{N}+Z\right)^2}\right) \end{aligned} \tag{57}$$

with $Z := \|\tilde{e}_t\|^2/N$. Since $\|\tilde{e}_t\|^2$ being a sum of N squared t-distributed random numbers with degrees-of-freedom-parameter ν, we know that $Z \sim F(N,\nu_e)$ (see, for example, p.22 in Fang et al. (1990)). Knowing the probability density function of Z, we can compute the desired expected value:

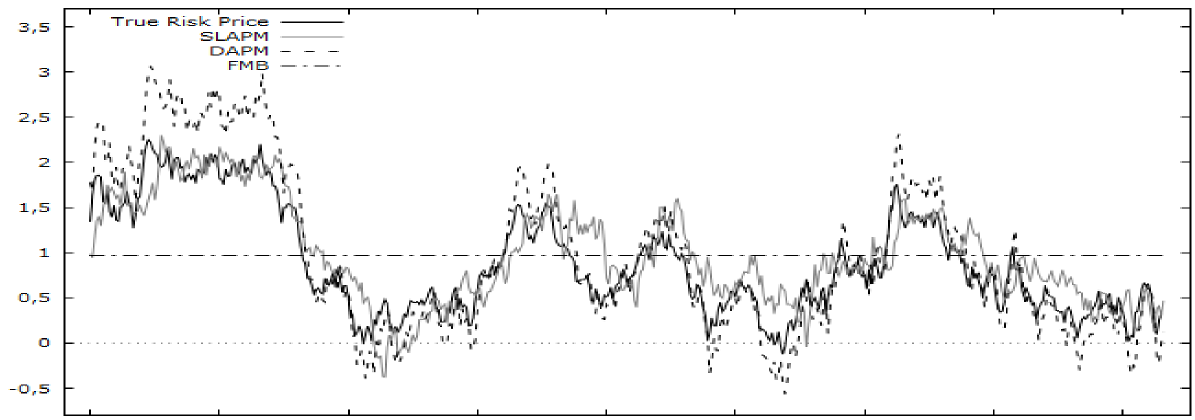
$$\begin{aligned} \mathbb{E}_{t-1}\left(\frac{Z}{\left(\frac{\nu_e}{N}+Z\right)^2}\right) &= \int_0^\infty \frac{z}{\left(\frac{\nu_e}{N}+z\right)^2}\frac{(N/\nu_e)^{\frac{N}{2}}}{B\left(\frac{N}{2},\frac{\nu_e}{2}\right)}z^{\frac{N}{2}-1}\left(1+\frac{N}{\nu_e}z\right)^{-\frac{N+\nu_e}{2}}dz \\ &= \frac{N/\nu_e}{B\left(\frac{N}{2},\frac{\nu_e}{2}\right)}\int_0^\infty\left(\frac{N}{\nu_e}z\right)^{\frac{N}{2}}\left(1+\frac{N}{\nu_e}z\right)^{-\frac{N+\nu_e}{2}-2}\frac{N}{\nu_e}dz \qquad \Big|\ \text{substitute with } t := \frac{N}{\nu_e}z \\ &= \frac{N/\nu_e}{B\left(\frac{N}{2},\frac{\nu_e}{2}\right)}\int_0^\infty t^{\frac{N+2}{2}-1}(1+t)^{-\frac{N+2}{2}-\frac{\nu_e+2}{2}}dt \\ &= \frac{N}{\nu_e}B\left(\frac{N}{2},\frac{\nu_e}{2}\right)^{-1}B\left(\frac{N+2}{2},\frac{\nu_e+2}{2}\right) \\ &= \frac{N}{\nu_e}B\left(\frac{N}{2},\frac{\nu_e}{2}\right)^{-1}B\left(\frac{N}{2},\frac{\nu_e+2}{2}\right)\frac{N/2}{N/2+\nu_e/2+1} \\ &= \frac{N}{\nu_e}B\left(\frac{N}{2},\frac{\nu_e}{2}\right)^{-1}B\left(\frac{N}{2},\frac{\nu_e}{2}\right)\frac{N/2}{N/2+\nu_e/2+1}\frac{\nu_e/2}{N/2+\nu_e/2} \\ &= \frac{N^2}{(N+\nu_e+2)(N+\nu_e)} \end{aligned} \tag{58}$$

where we used the integral representation of the beta function $B(x,y) = \int_0^\infty t^{x-1}(1+t)^{-x-y}dt$ in the fourth equation and the identities $B(x+1,y) = B(x,y)\cdot x/(x+y)$ and $B(x,y+1) = B(x,y)\cdot y/(x+y)$ in the fifth and sixth equations respectively.
Inserting (58) into (57) yields $C_{\mathcal{I}}(\tilde{e}_t,\psi^e) = \frac{\nu_e+N}{\nu_e+N+2}$ and therefore proves the assertion in conjunction with Proposition 1 and $\mathbb{E}\left(u_tu_t^\top\right) = \frac{\nu_u}{\nu_u-2}\Omega_u$.

□



(a) Exemplary Simulated Return Series



(b) Exemplary Simulated Risk Price Series

Figure 1: **Exemplary simulated Return and Lambda.** The figure shows simulated excess returns together with the conditional expectation $\beta\lambda_{t-1}$(panel (a)) and the associated lambda series (panel (b)). The simulated cross-section consists of $N = 25$ assets and $T = 500$ observations. The considered lambda forecasters in panel (b) are a Gaussian constant beta SLAPM(1,1) specification (SLAPM), a constant risk price specification estimated fitted with the classical approach of Fama and MacBeth (1973) (FMB) and the regression-based dynamic asset pricing model of Adrian et al. (2015) (DAPM) with the correct forecasting factor ($\kappa = 1$).

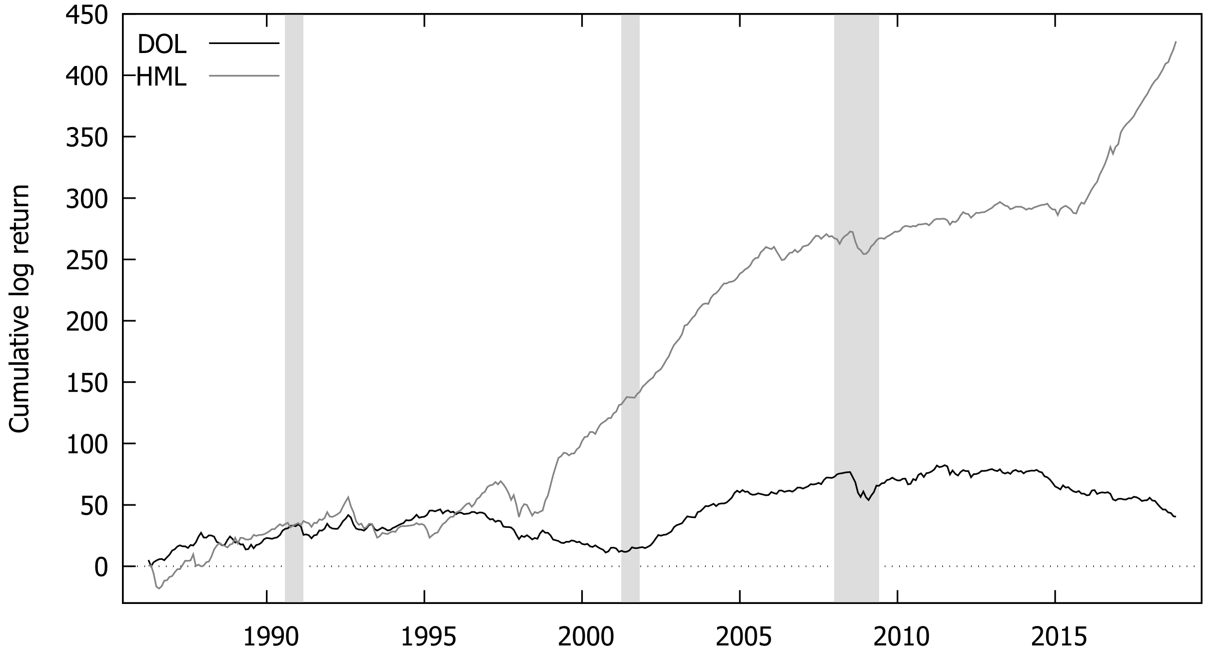
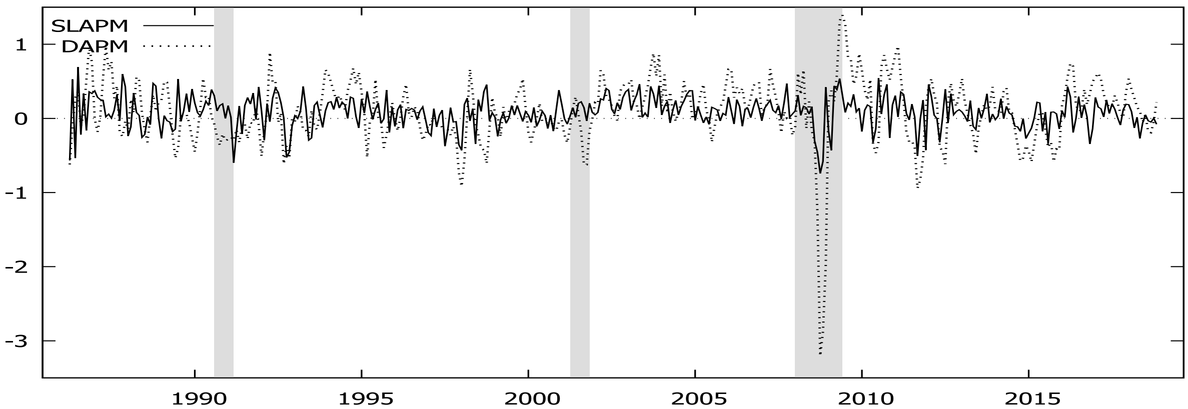
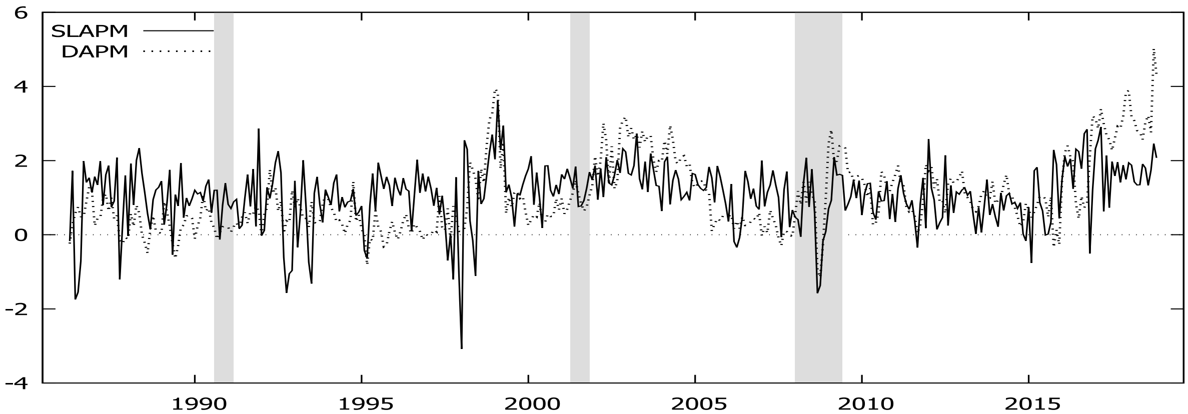


Figure 2: **Cumulative DOL and HML Log Returns.** The figure shows time-varying cumulative log returns to the DOL and HML portfolio. DOL is a equally-weighted portfolio of every currency available and HML is a carry trade portfolio with a long position in a interest rate currency portfolio and a short position in a low interest rate portfolio. Shaded areas refer to NBER recessions. The sample period is 1986:04 - 2018:11.

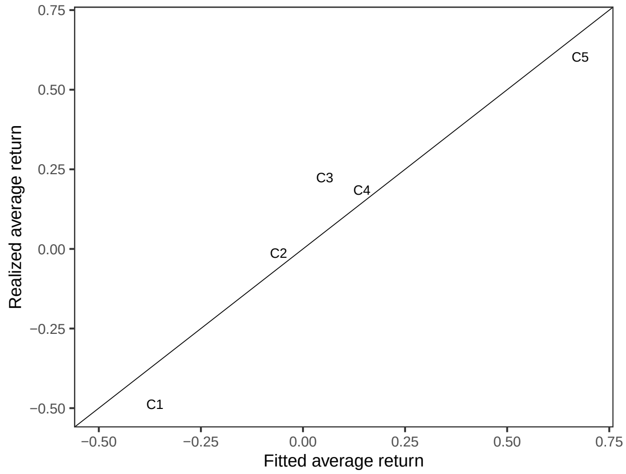


(a) DOL Risk Premium

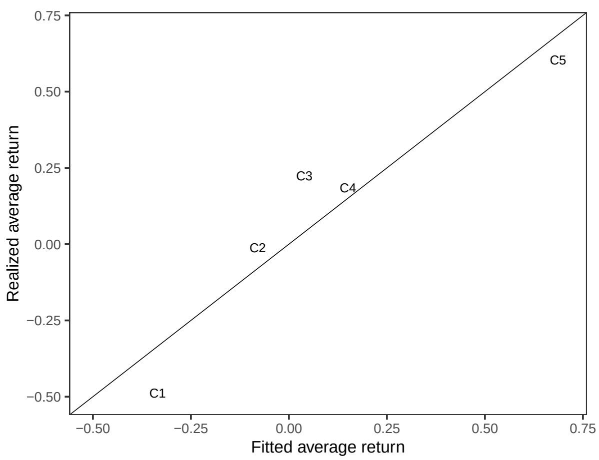


(b) HML Risk Premium

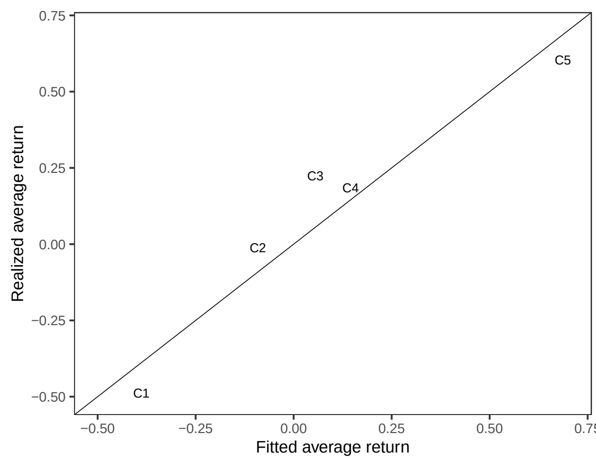
Figure 3: **Risk Price Dynamics in the Carry Cross-Section.** The figure shows time-varying risk premia of the DOL (rp_t^{DOL}) as well as the HML (rp_t^{HML}) factor from the two dynamic estimation approaches. SLAPM refers to results from a Gaussian SLAPM(1,1) model with constant risk exposures. DAPM refers to results from the regression-based approach of Adrian et al. (2015) using forecast variables discussed in Section 4.1.2. Test assets are five currency portfolios sorted on forward discount. The sample period is 1986:04 - 2018:11. Shaded areas refer to NBER recessions.



(a) Unconditional Model



(b) DAPM



(c) SLAPM

Figure 4: **Scatter Plots Carry Cross-Section.** This figure shows plots of observed against model-implied average returns estimated using the three different estimation approaches. SLAPM refers to results from a Gaussian SLAPM(1,1) model with constant risk exposures. DAPM refers to results from the regression-based approach of Adrian et al. (2015) using forecast variables discussed in Section 4.1.2. Unconditional model estimation is conducted with a DAPM without forecast variables. Test assets are five currency portfolios sorted on forward discount. The sample period is 1986:04 - 2018:11.

Table 1: Risk Price Forecast Simulation OOS-R^2

The table shows the OOS-R^2 of Gaussian constant beta SLAPM(1,1) risk price forecasts against competing benchmark models. The eight simulated panels have different numbers of assets N, time observations T and are replicated 10 000 times each. Benchmarks are a constant risk price specification estimated fitted with the classical approach of Fama and MacBeth (1973) and the regression-based dynamic asset pricing model of Adrian et al. (2015) (DAPM). The share of information from the correct forecaster made available to the DAPM is denoted with κ. Shaded areas indicate that the Diebold and Mariano (1995) test with null hypothesis of equal forecast accuracy can not be rejected.

		N=25			N=50			N=100		
		T= 250	500	1000	T=250	500	1000	T=250	500	1000
FMB		0.587	0.699	0.775	0.594	0.692	0.741	0.652	0.686	0.724
DAPM	$\kappa = 0$	0.664	0.738	0.790	0.676	0.730	0.759	0.723	0.719	0.744
	0.1	0.661	0.734	0.787	0.674	0.727	0.756	0.720	0.717	0.741
	0.2	0.650	0.723	0.777	0.663	0.716	0.746	0.711	0.708	0.728
	0.3	0.625	0.697	0.755	0.640	0.691	0.721	0.691	0.684	0.699
	0.4	0.582	0.648	0.708	0.599	0.642	0.670	0.657	0.635	0.639
	0.5	0.518	0.563	0.620	0.537	0.559	0.574	0.607	0.549	0.527
	0.6	0.440	0.433	0.464	0.460	0.435	0.402	0.547	0.414	0.330
	0.7	0.369	0.273	0.217	0.386	0.281	0.126	0.490	0.246	0.021
	0.8	0.319	0.124	-0.086	0.331	0.138	-0.230	0.449	0.091	-0.359
	0.9	0.294	0.033	-0.321	0.301	0.046	-0.532	0.426	0.000	-0.661
	1	0.288	0.010	-0.386	0.293	0.018	-0.642	0.419	-0.020	-0.758

Table 2: Return Forecast Simulation OOS-R^2(%)

The table shows the percentage OOS-R^2 of Gaussian constant beta SLAPM(1,1) return forecasts against competing benchmark models computed from simulations. The eight simulated panels have different numbers of assets N, time observations T and are replicated 10 000 times each. Benchmarks are sample mean forecasts (Mean) and forecasts from predictive regressions (PReg) with a diffuse signal of the correct forecaster as regressor. κ is the share of information from the correct forecaster. Shaded areas indicate that the Diebold and Mariano (1995) test with null hypothesis of equal forecast accuracy can not be rejected.

		N=25			N=50			N=100		
		T= 250	500	1000	T=250	500	1000	T=250	500	1000
(a) $\beta = 0.5$										
Mean		0.729	0.513	0.807	0.690	0.830	0.523	0.563	0.802	0.699
PReg	$\kappa = 0$	1.053	0.764	0.968	1.115	1.044	0.531	0.971	0.954	0.830
	0.1	1.106	0.755	0.979	1.127	1.088	0.520	0.991	0.951	0.842
	0.2	1.143	0.716	0.962	1.116	1.109	0.493	1.006	0.926	0.824
	0.3	1.146	0.632	0.903	1.072	1.086	0.445	1.011	0.869	0.757
	0.4	1.100	0.501	0.794	0.992	1.004	0.377	1.000	0.778	0.630
	0.5	1.004	0.339	0.644	0.885	0.862	0.297	0.975	0.664	0.453
	0.6	0.881	0.184	0.484	0.772	0.689	0.227	0.948	0.551	0.263
	0.7	0.758	0.065	0.347	0.675	0.524	0.178	0.928	0.460	0.102
	0.8	0.652	-0.007	0.249	0.601	0.390	0.153	0.914	0.398	-0.012
	0.9	0.570	-0.042	0.188	0.551	0.293	0.146	0.903	0.363	-0.079
	1	0.509	-0.051	0.156	0.521	0.229	0.150	0.894	0.348	-0.112
(b) $\beta = 1$										
Mean		0.877	0.671	1.064	0.628	0.772	0.881	0.714	0.857	1.042
PReg	$\kappa = 0$	1.355	0.936	1.209	1.006	1.044	0.925	1.198	0.998	1.184
	0.1	1.411	0.955	1.196	1.037	1.055	0.888	1.192	0.979	1.185
	0.2	1.435	0.941	1.144	1.052	1.032	0.813	1.165	0.926	1.141
	0.3	1.405	0.874	1.034	1.042	0.958	0.685	1.105	0.827	1.031
	0.4	1.304	0.743	0.862	1.002	0.824	0.507	1.013	0.684	0.842
	0.5	1.143	0.563	0.651	0.937	0.643	0.304	0.907	0.519	0.595
	0.6	0.961	0.373	0.444	0.862	0.454	0.121	0.815	0.368	0.343
	0.7	0.793	0.212	0.282	0.789	0.297	-0.007	0.755	0.256	0.137
	0.8	0.658	0.098	0.179	0.725	0.187	-0.073	0.725	0.189	-0.001
	0.9	0.557	0.027	0.128	0.673	0.122	-0.090	0.716	0.158	-0.077
	1	0.486	-0.013	0.111	0.632	0.089	-0.078	0.716	0.150	-0.108
(c) $\beta = 1.5$										
Mean		0.993	0.593	1.126	0.799	0.903	0.907	0.904	0.814	1.058
PReg	$\kappa = 0$	1.451	0.768	1.289	1.169	1.062	0.884	1.342	0.968	1.203
	0.1	1.503	0.783	1.277	1.166	1.086	0.856	1.353	0.937	1.215
	0.2	1.521	0.759	1.220	1.135	1.077	0.791	1.339	0.869	1.186
	0.3	1.481	0.675	1.102	1.066	1.013	0.679	1.283	0.755	1.094
	0.4	1.368	0.521	0.916	0.962	0.883	0.519	1.181	0.598	0.925
	0.5	1.193	0.313	0.686	0.841	0.696	0.334	1.044	0.428	0.694
	0.6	0.993	0.099	0.462	0.733	0.493	0.165	0.907	0.282	0.449
	0.7	0.810	-0.079	0.286	0.653	0.318	0.044	0.796	0.185	0.243
	0.8	0.665	-0.202	0.173	0.602	0.190	-0.020	0.721	0.135	0.099
	0.9	0.559	-0.275	0.116	0.572	0.108	-0.040	0.676	0.120	0.015
	1	0.487	-0.312	0.097	0.556	0.062	-0.033	0.652	0.125	-0.026

Table 3: Price of Risk Estimates

This table shows estimates of risk price parameters for the risk factors DOL and HML. The first two columns show results from a Gaussian constant-beta SLAPM(1,1) with score coefficient a and autoregressive coefficient b. Following four rows provide estimates from a dynamic asset pricing model in line with Adrian et al. (2015). Forecasting factors a5re the absolute average forwars discount as well as 3-monthly log growth rates of FX volatility, the CRB commodity return, the aggregated intermediary capital ratio of He et al. (2017) and the FRB dollar index. The last two columns show results from a constant risk price specification estimated with Fama-MacBeth regressions. Standard errors are shown in parentheses. They are computed with GMM for the Fama-MacBeth regression results and according to Adrian et al. (2015) for the the two DAPM specifications. They both adjust for cross-asset correlation in the residuals and for estimation error of the time-series betas. SLAPM standard errors are derived from the numerically computed Fisher Information as described in Section 2.3. Test assets are 5 equally-weighted currency portfolios sorted on forward discount with monthly returns denoted in percentages. The sample period is 1986:04 - 2018:11.

	SLAPM		DAPM		Unconditional	
	DOL	HML	DOL	HML	DOL	HML
const	0.075***	0.872***	-0.111	-0.033	0.091	1.042***
	(0.026)	(0.108)	(0.153)	(0.212)	(0.102)	(0.147)
a	0.836***	1.132***				
	(0.072)	(0.078)				
b	0.208***	0.356***				
	(0.061)	(0.052)				
ΔFX Volatility			-0.027***	-0.029**		
			0.010	0.015		
ΔCommodity Returns			0.305***	0.159*		
			(0.058)	(0.081)		
Abs. Avg. Forward Discount			0.013	0.142***		
			(0.015)	(0.021)		
ΔIntermediary Capital			-0.103***	-0.043		
			(0.026)	(0.037)		
ΔBroad Dollar Index			0.298**	-0.064		
			(0.124)	(0.172)		

*p<0.1; **p<0.05; ***p<0.01

Table 4: Mean Squared (Forecast) Error Comparisons

This table compares root mean squared errors and root mean squared forecast error across the three estimation approaches of the currency cross-section model. SLAPM refers to results from a Gaussian SLAPM(1,1) model with constant risk exposures. The in-sample version employs lambdas updated at the current observation. DAPM refers to results from the regression-based approach of Adrian et al. (2015) using forecast variables discussed in Section 4.1.2. Unconditional model estimation is conducted with a DAPM without forecast variables. Test assets are five currency portfolios sorted on forward discount. HML and DOL refer to errors from predicting the equally-weighted currency market return and returns to the high-minus-low carry trade return, respectively. The sample period is 1986:04 - 2018:11.

	C1	C2	C3	C4	C5	Avg	DOL	HML
(a) Root Mean Squared Error								
Unconditional	0.889	0.878	0.879	0.932	0.949	0.905	0.269	0.937
DAPM	0.824	0.887	0.852	0.900	0.782	0.849	0.144	0.599
SLAPM	0.781	0.915	0.895	0.944	0.866	0.880	0.102	0.702
SLAPM (in-sample)	0.827	0.867	0.862	0.904	0.785	0.849	0.236	0.566
(b) Root Mean Squared Forecast Error								
Unconditional	2.359	1.990	2.228	2.316	2.812	2.341	2.006	2.845
DAPM	2.244	1.930	2.149	2.219	2.652	2.239	1.909	2.620
SLAPM	2.297	1.992	2.225	2.310	2.756	2.316	1.994	2.700
SLAPM (in-sample)	2.153	1.879	2.088	2.157	2.388	2.133	1.852	2.131