Moment Conditions and Time-Varying Risk Premia

Dennis Umlandt*

University of Innsbruck

Abstract

This paper proposes a novel approach for estimating linear factor pricing models with dynamic risk premia based on a generalized method of moments (GMM) framework. Time-varying risk prices and exposures follow an updating scheme that aims for the steepest descent of the conditional moment-criterion function. The most informative moment for inferring risk premium dynamics comes from the cross-sectional pricing equation estimated in the second stage of the widely used Fama-MacBeth regression approach. Monte Carlo results show that the new approach is able to adequately filter various types of risk premium dynamics. An application to the Fama-French 5-factor model shows that the GMM-based procedure can largely reduce pricing errors compared to other dynamic and static approaches. The results show that premium dynamics vary across factors, and while they are generally countercyclical, they exhibit significant declines at the beginning of crisis periods.

JEL Codes: G12, G17, C58

*The author thanks Heiner Beckmeyer (discussant), Jeroen Dalderop (discussant), Moritz Dauber, Giorgia Simion (discussant), Sebastian Stöckl, seminar participants at the Universities of Innsbruck, Trier, Liechtenstein and Graz as well as participants of the Statistical Week 2023, the 38th workshop of the Austrian Working Group in Banking and Finance (AWG) 2023, the German Finance Association (DGF) Annual Meeting 2023, the 16th International Conference on Computational and Financial Econometrics (CFE) 2023, and the Midwest Finance Association (MFA) Annual Meeting 2024 for valuable discussions and comments.

1 Introduction

Financial theories interpret expected returns over a risk-free interest rate, known as excess returns, as compensation for the risk of the investment. Traditional factor asset pricing models (Fama and French, 1993; 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. Estimation of these linear factor models, which are widely used in empirical financial research, are typically conducted with the traditional two-step regression approach of Fama and MacBeth (1973) (henceforth denoted with FMB) or Generalized Method of Moments (GMM) frameworks following Hansen (1982). However, despite crucial evidence in the literature that risk premia vary over time (Campbell and Shiller, 1988; Fama and French, 1989; Cochrane, 2011), risk exposures of financial securities, as well as risk prices and thus risk premia, are typically assumed to be constant over time in these estimation approaches. Moreover, the recent factor timing literature (Moreira and Muir, 2017; Haddad et al., 2020; Ehsani and Linnainmaa, 2022; Arnott et al., 2023; Neuhierl et al., 2023) has shown that the dynamics of risk premia can explain economically relevant excess returns and are therefore important for understanding the factor structure of financial returns.

This paper proposes a moment-based approach to estimate time-varying risk premia in linear factor pricing models. The approach builds on a small set of asset pricing moments which is often used for GMM-based estimation of unconditional asset pricing models following the methodology of Hansen (1982). We extend this baseline model with an observation-driven updating scheme for risk exposures and prices in order to achieve a conditional factor pricing model. The parameter updating follows the general approach of Creal et al. (2024) to let the dynamics be driven by the influence function of the conditional GMM estimator in each time period. This intuitively provides a steepest descent improvement of the local GMM criterion function in the corresponding time period. It turns out that a such constructed updating mechanism adjusts risk prices according to regression errors from the cross-sectional regression performed in the second stage of the FMB procedure. Thus, instead of finding a risk price that minimizes these errors on average, the procedure here uses these cross-sectional pricing errors to infer risk price dynamics. If the idiosyncratic innovation of an asset moves with a factor innovation, the factor exposure of that asset is increased by raising the corresponding beta to remove the unwanted comovement. Static parameters in the introduced Moment-Based Dynamic Asset Pricing Model (MDAPM) can be readily estimated using an instrumented GMM approach. As usual for GMM, more efficient estimates can be obtained by performing moment minimization with the optimal weighting matrix in a second stage.

The performance of the MDAPM is evaluated with a Monte Carlo study. In a scenario with a realistically low signal-to-noise ratio, the MDAPM can recover various risk premium dynamics such as cycles and structural breaks. Compared to a static benchmark, the new dynamic approach successfully reduces both pricing and risk premium prediction errors.

We apply the GMM-based dynamic framework to the Fama-French 5-factor model (Fama and French, 2015) on a cross-section of characteristics-sorted equity portfolios. It can substantially reduce pricing and risk premium prediction errors compared to an unconditional model and the regression-based dynamic asset pricing model of Adrian et al. (2015), which uses stock return predictors to infer risk price movements. The results suggest that in the cross-section of 32 portfolios sorted by market capitalization, operating profitability, and investment, the variation in the risk premium is more strongly driven by changes in risk prices than by changes in betas.

We document three main observations with respect to filtered risk premia. First, risk premia initially fall at the beginning of recessions for several months but rise afterwards. A pattern that have also been documented by Gómez-Cram (2022) for the market premium. Thus, it appears that after an initial downward adjustment of expectations, adverse events lead investors to demand higher compensation for risk. This is puzzling given the numerous references to the countercyclical behavior of expected excess returns on stocks, such as Fama and French (1989), Ferson and Harvey (1991), and Lustig and Verdelhan (2012). Second, filtered risk prices are particularly consistent with the trajectories implied by stock return predictors in the case of the market risk premium, while there are substantial discrepancies for the other factor premia such as value or operating profitability. Therefore, predictors that predict overall market returns may not be adequate or sufficient instruments for predicting other factor risk premia. The third major observation is that based on risk premium prediction errors and in line with the recent literature on factor momentum (Arnott et al., 2023; Ehsani and Linnainmaa, 2022), the momentum premium can be explained by variation in the risk prices of other factors. However, we also find that pricing errors can be substantially reduced by including a cross-sectional momentum factor, suggesting that momentum exposure is only partially priced and that this priced exposure can be explained by the momentum of the other factors.

The search for empirical methods to estimate factor asset pricing models with time-varying risk premia has recently received renewed attention. Regression-based approaches such as Adrian et al. (2015, 2019), Gagliardini et al. (2016, 2020) and Chaieb et al. (2021) use instrument variables to explain the time dynamics of the parameters λ and β. A common drawback of these approaches is that time-varying risk premia can only be identified with respect to a filtration spanned by the set of instruments employed. This is particularly problematic given that the literature on the predictability of returns is still debating whether returns are predictable at all and what the appropriate predictors are. Thus, even if the literature finds that risk premia are significantly time-varying, the interpretation of the filtered premia series may be misleading due to inappropriate or simply missing predictors. The GMM-based dynamic model proposed here avoids this problem by filtering out the dynamics of risk premia from the full set of available assets and factors. The time series predictors used in the regression-based approaches can additionally be used as instruments to identify factor mean dynamics.

Umlandt (2023) and Giroux et al. (2024) provide observation-driven filters to estimate dynamic financial risk premia without the need to specify instrument variables for time dynamics. These likelihood-based methods follow the generalized autoregressive score approach of Creal et al. (2013). Although a misspecification bias due to inappropriate time series predictors can be circumvented, explicit distributional assumptions are required. In contrast, the asset pricing literature often refrains from posing distributional assumptions and specifies a set of moment restrictions instead that are typically derived from no-arbitrage assumptions. The GMM-based approach proposed here develops a filter for risk premia based only on such a set of moment restrictions, but otherwise follows the logic of the likelihood-based filter. Thus, our approach can be seen as an alternative that is robust to distributional misspecification. Another advantage of our approach is that, in contrast to the likelihood-based model, we do not have to explicitly estimate the covariance matrix of the asset-specific innovations. This massively reduces the number of parameters in the numerical optimization and thus lowers the computational burden of the estimation procedure considerably.

Time-varying risk premia are also central features of machine learning-inspired methods such as the instrumental principal component analysis (IPCA) of Kelly et al. (2019) and the risk premium principal component analysis (RP-PCA) of Lettau and Pelger (2020). These methods assume that cross-sectional pricing factors are latent and estimate them together with time-varying risk exposures. In contrast, the method proposed in this paper and those mentioned above assume that cross-sectional factors are observed and focus on inferring the time dynamics of risk prices and exposures.

The remainder of the paper is organized as follows. Section 2 introduces and discusses the dynamic GMM framework for linear factor pricing models. The empirical application on the Fama-French 5-Factor model is presented in Section 3. Section 4 concludes.

2 Dynamic GMM Model

In the following we introduce the dynamic GMM-model upon a fairly general baseline factor pricing model that in a similar fashion serves as basis for most of the employed empirical methods, for example, Fama and MacBeth (1973), Ang and Kristensen (2012), Adrian et al. (2015, 2019), Umlandt (2023), and Gagliardini et al. (2016, 2020).

2.1 Baseline Model

Let $r_t = (r_t^1, \ldots, r_t^N)^\top$ denote the N-dimensional vector representing the excess returns of N different assets at time $t \in \{0, \ldots, T\}$. The underlying data-generating process is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with a filtration $\mathcal{F}_t = \sigma(\{r_t, \ldots, r_0\})$ 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

$$f_t = \phi_0 + \phi z_{t-1} + u_t, \qquad t = 1, \ldots, T, \tag{1}$$

where u_t is an independently and identically distributed zero mean noise term with covariance matrix Σ_u ($u_t \stackrel{iid}{\sim} (0, \Sigma_u)$), and z_t is a L-dimensional vector of lagged predictors adapted to $\mathcal{F}_t$. Predictors may be external as well as past factor observations. We refer to equation (1) as the risk factor model.

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

$$\mathbb{E}_{t-1}(m_t r_t) = 0, \tag{2}$$

where $\mathbb{E}_{t-1}$ denotes the conditional expectation with respect to time $t-1$ information $\mathcal{F}_{t-1}$. The Euler equation (2) can be used to compute the conditional covariance between the SDF and the asset return as

$$Cov_{t-1}(r_t, m_t) = -\mathbb{E}_{t-1}(r_t)\mathbb{E}_{t-1}(m_t). \tag{3}$$

Regressing the demeaned returns on the factor innovations u_t yields an N-dimensional idiosyncratic noise term $e_t \stackrel{iid}{\sim} (0, \Sigma_e)$ that is orthogonal to u_t. Taken together with (3), the return can be decomposed as

$$r_t = \mathbb{E}_{t-1}(r_t) + (r_t - \mathbb{E}_{t-1}(r_t)) = -\frac{Cov_{t-1}(r_t, m_t)}{\mathbb{E}_{t-1}(m_t)} + \beta_{t-1}u_t + e_t, \tag{4}$$

where $\beta_{t-1} = Cov_{t-1}(r_t, u_t)\Sigma_u^{-1}$ denotes the $N \times K$-dimensional matrix of risk exposures.

In order to transform equation (4) into a cross-sectional pricing model, assume the SDF to be affine-linear in the economy's risk factor innovations; that is,

$$\frac{m_t - \mathbb{E}_{t-1}(m_t)}{\mathbb{E}_{t-1}(m_t)} = -\lambda_{t-1}^{\top}\Sigma_u^{-1}u_t \tag{5}$$

with time-varying price of risk vector λ_{t-1} of dimension K. Plugging the SDF into the return decomposition (4) yields a standard beta representation given by

$$r_t = Cov_{t-1}\left(r_t, u_t\right)\Sigma_u^{-1}\lambda_{t-1} + \beta_{t-1}u_t + e_t \tag{6}$$

$$= \beta_{t-1}\lambda_{t-1} + \beta_{t-1}u_t + e_t. \tag{7}$$

The return decomposition (7) therefore consists of a predictable risk premium $\beta_{t-1}\lambda_{t-1}$ that compensates risk exposures, an unpredictable component $\beta_{t-1}u_t$ depending on risk factor innovations and an asset-specific innovation term e_t. Representation (7) is also referred to as the cross-sectional pricing equation.

2.2 Moment Conditions

Given constant risk premia, that means $\beta_t \equiv \beta$ and $\lambda_t \equiv \lambda$, the model in Section 2.1 can be estimated with GMM using a concise set of moment conditions. The static model moment conditions introduced here will serve as the basis for the dynamic model that will be derived in the following sections.

The first two sets of moment conditions stem from the factor model equation (1) and are given by

$$\mathbb{E}_{t-1}\left[u_t\right] = \mathbb{E}_{t-1}\left[f_t - \phi_0 - \phi z_{t-1}\right] = 0 \tag{8}$$

$$\mathbb{E}_{t-1}\left[\operatorname{vec}\left(u_t z_{t-1}^{\top}\right)\right] = \mathbb{E}_{t-1}\left[\operatorname{vec}\left(\left(f_t - \phi_0 - \phi z_{t-1}\right) z_{t-1}^{\top}\right)\right] = 0 \tag{9}$$

where vec() refers to the vectorization operator that stacks all columns of a matrix on top of one another into one column vector. Equation (8) states that the factor innovation should be zero on average, whereas equation (9) is a standard orthogonality assumption that requires the set of instruments z_{t-1} to be uncorrelated with the factor innovations u_t.

Furthermore, we use the aforementioned orthogonality assumption between factor innovations and idiosyncratic innovations to set up the third moment condition given by

$$\mathbb{E}_{t-1}\left[\operatorname{vec}\left(e_t u_t^{\top}\right)\right] = \mathbb{E}_{t-1}\left[\operatorname{vec}\left((r_t - \beta\lambda - \beta u_t) u_t^{\top}\right)\right] = 0. \tag{10}$$

According to equation (10), the conditional covariance between the return vector and factor innovations must satisfy $\mathbb{E}_{t-1}\left[r_t u_t^{\top}\right] = \beta\Sigma_u$. Thus, the third moment condition (10) identifies betas as the time series regression coefficients whose least squares estimators are used in the traditional regression approach of Fama and MacBeth (1973).

The final fourth set of conditions sets the conditional expectations of cross-sectional pricing errors to zero, i.e.

$$\mathbb{E}_{t-1}\left[e_t\right] = \mathbb{E}_{t-1}\left[r_t - \beta\lambda - \beta u_t\right] = 0. \tag{11}$$

Note that conditions (8) and (11) identify the risk premium given by $\mathbb{E}_{t-1}\left(r_t\right) = \beta\lambda$ which is of major interest in the following. We stack the $M = K(N + L + 1) + N$ conditions (8) to

(11) as conditional moment function given by

$$g_t(x_t;\theta_0) = \begin{pmatrix} u_t \\ \mathrm{vec}\left(u_t z_{t-1}^\top\right) \\ \mathrm{vec}\left(e_t u_t^\top\right) \\ e_t \end{pmatrix} \tag{12}$$

with observation vector $x_t = (r_t, f_t, z_{t-1})$ and an $K(N+L+1)+K$-dimensional parameter vector $\theta_0^\top = (\phi_0^\top, \mathrm{vec}(\phi)^\top, \mathrm{vec}(\beta)^\top, \lambda^\top)$. Note that in a typical asset pricing application, the number of test assets, N, exceeds the number of factors, K. Therefore, the model is generally over-identified because the number of moment conditions exceeds the number of parameters. In order to just identify the parameters and give particular weight to the pricing model condition (11), we weight the moment conditions with the $K \times N$ matrix $\beta^\top$. This weighting also identifies λ with its cross-sectional OLS estimate.

The conditional GMM criterion function of the asset pricing model to be minimized for estimation can then be written as

$$\mathbb{E}_{t-1}\left[g_t(x_t;\theta_0)\right]^\top \Omega \mathbb{E}_{t-1}\left[g_t(x_t;\theta_0)\right] \tag{13}$$

with weighting matrix

$$\Omega = \begin{pmatrix} I_{M-N} & 0_{(M-N)\times N} \\ 0_{N\times(M-N)} & \beta\beta^\top \end{pmatrix}. \tag{14}$$

2.3 Time-Varying Risk Premia

In the following, we want to uncover dynamic risk premia by extending the static baseline model of Section 2.1 by an observation-driven updating scheme for risk premium parameters given by

$$\vartheta_t = \overline{\vartheta} + A s_{t-1} + B(\vartheta_{t-1} - \overline{\vartheta}). \tag{15}$$

where A and B are parameter matrices of appropriate size, and ϑ_t is a vector containing parameters from θ_0 which are supposed to vary over time. The parameter vector $\overline{\vartheta}$ represents the unconditional values of the time-varying parameters. Typically, we will choose either $\vartheta_t = \lambda_{t-1}$ or $\vartheta_t = \left(\lambda_{t-1}^\top, vec(\beta_{t-1})^\top\right)^\top$ in order to introduce dynamic risk premia.[1]

A crucial modeling decision that must be made to use the updating scheme in (15) is the specification of s_t. This quantity is intended to provide information about the direction in which the parameters should be updated, taking into account the information at time t, represented by the most recent observations r_t and f_t. The score-driven model class of Creal et al. (2013) and Harvey (2013) employs such a scheme, using the gradient of log observation density as an innovation sequence s_t for the dynamic adjustment. In particular, Umlandt (2023) studies a score-driven model using an asset pricing framework closely related to the one in Section 2.1, which additionally needs to specify the distribution of the innovation terms u_t and e_t.

Here, we want to avoid distributional assumptions about the error terms and let the dynamics be guided only by the moment conditions discussed in Section 2.2. We therefore follow the approach of Creal et al. (2024) and choose the innovation sequence s_t as the influence function of x_t on the estimator of ϑ_t from the conditional moment function (12). Let Δ_{x_t} be the Dirac measure that puts unit mass on the actual observation x_t. Given $\epsilon \in [0, 1]$, define the contaminated measure $F_x^\epsilon = (1 - \epsilon)F_x + \epsilon\Delta_{x_t}$ that overweights the current observation relative to the overall observational measure F_x. GMM estimates of the time-varying parameters can then be derived based on the overall measure as $\vartheta_t(F_x)$ or based

[1]The discrepancy in the time index between ϑ_t and, for instance, λ_{t-1} is simply due to different national conventions. While the asset pricing literature would denote the risk price as λ_{t-1}, since it is pricing returns in period t but is already known in $t-1$, the observation-driven model literature would usually denote such a predetermined parameter with λ_t or $\lambda_{t|t-1}$.

on the contaminated measure $\vartheta_t(F_x^\varepsilon)$. The influence function is then defined as the limit of the (functional) difference quotient of the two estimators as ϵ tends to 0, i.e.

$$s_t = \left.\frac{d\vartheta_t(F_x)}{d\epsilon}\right|_{\epsilon=0} = \lim_{\epsilon\to 0} \frac{\vartheta_t\left(F_x^\epsilon\right) - \vartheta_t\left(F_x\right)}{\epsilon}. \tag{16}$$

Intuitively, the influence function measures the dependence of (static) parameter estimators on new observations. This concept is widely used in robust statistics to study the impact of data outliers on estimates.[2] But this also means that the influence function provides a signal of where one can get the steepest descent in the conditional GMM criterion (12) in response to the new observation x_t. Thus, instead of focusing on outliers, we use the information from the influence function to guide the updating of the time-varying risk premium parameter.

Let $\tilde{\theta}_0$ include the static parameters from θ_0 which are not covered in ϑ_t and $g_t(x_t; \vartheta_t, \tilde{\theta}_0)$ be the conditional moment criterion in (12) but including time-varying parameters ϑ_t. Creal et al. (2024) show that choosing s_t as the influence function in a restricted observation-driven updating scheme like (15) delivers a local expected improvement of the conditional criterion function given by

$$\mathbb{E}_{t-1}\left[g_t(x_t; \vartheta_t, \tilde{\theta}_0)\right]^\top \Omega_{t-1} \mathbb{E}_{t-1}\left[g_t(x_t; \vartheta_t, \tilde{\theta}_0)\right] \tag{17}$$

where Ω_{t-1} is a weighting matrix known at time $t-1$. We choose the dynamic weighting matrix to be either the one given in (14) or, if betas are included in ϑ as

$$\Omega_{t-1} = \begin{pmatrix} I_{M-N} & 0_{(M-N)\times N} \\ 0_{N\times(M-N)} & \beta_t \beta_t^\top \end{pmatrix}. \tag{18}$$

Note that in the latter specification, Ω_{t-1} is still known at time $t-1$, since $\beta_t = Cov_{t-1}(r_t, u_t)\Sigma_u^{-1}$

[2]See Hampel et al. (2011) for a comprehensive treatment of the influence function and its use in robust statistics.

is known at time $t-1$.

We say a common series $x_t^\top = \left(r_t^\top, f_t^\top, z_{t-1}^\top\right)$ of returns r_t, (cross-sectional) factors f_t, and time-series predictors z_{t-1} follows a **M**oment-Based **D**ynamic **A**sset **P**ricing **M**odel (MDAPM) if it fulfills the moment conditions (8) to (11) and includes an updating scheme given by (15) and (16). The following proposition presents the updating schemes for the cases in which either only risk prices or both risk prices and exposures vary over time.

Proposition 1. Let x_t follow an MDAPM process.

(a) The influence function for $\vartheta_t = \lambda_{t-1}$ is given by

$$s_t^\lambda = \left(\beta_{t-1}^\top \beta_{t-1}\right)^{-1} \beta_{t-1}^\top e_t \tag{19}$$

$$= \left(\beta^\top \beta\right)^{-1} \beta^\top r_t - \lambda_{t-1} - u_t \tag{20}$$

(b) The influence function for $\vartheta_t = \left(\lambda_{t-1}^\top, vec(\beta_{t-1})^\top\right)^\top$ is given by

$$\begin{pmatrix} s_t^\lambda \\ s_t^\beta \end{pmatrix} = \begin{pmatrix} \left(\beta_{t-1}^\top \beta_{t-1}\right)^{-1} \beta_{t-1}^\top e_t \left(1 - u_t^\top \Sigma_u^{-1} \lambda_{t-1}\right) \\ \text{vec}\left(e_t u_t^\top \Sigma_u^{-1}\right) \end{pmatrix} \tag{21}$$

The risk price updating scheme for the case with constant betas in (Proposition 1.(a)) resembles the regression error from a (cross-sectional) regression of r_t on the risk exposures β which in the given model should provide an estimate of $\lambda_{t-1} + u_t$ because of the asset pricing restriction (11). This cross-sectional regression is also employed in the second stage of the famous two-pass approach of Fama and MacBeth (1973) who compute period-wise lambdas from those regressions and average them to achieve an estimate of the (static) risk price. Instead of averaging the regression parameters, the MDAPM quite intuitively updates time-varying risk prices based on current pricing errors. A very similar updating scheme can also be found in the SDAPM of Umlandt (2023) in which the regression follows a generalized least squares fashion with $s_t^\lambda = \left(\beta^\top \Sigma_e^{-1} \beta\right)^{-1} \beta^\top \Sigma_e^{-1} e_t$ when assuming Gaussian innovations.

A major obstacle to the implementation of the Gaussian SDAPM is the appearance of the typically large covariance matrix Σ_e, which renders the procedure computationally demanding. Thus, the moment-based updating scheme is on the one hand much more convenient to implement, but on the other hand may lack efficiency when heteroskedasticity and correlation in idiosyncratic errors is a prevalent feature of the data. One way to utilize this potential correlation structure is to set the lower right component of the weighting matrix in (14) to $\Sigma_e^{-1}\beta\,(\Sigma_e^{-1}\beta)^\top$. Weighting the risk exposures with the error covariance matrix within the moment-based framework results in the SDAPM s_t^λ updating scheme.

Allowing risk prices and exposures to vary simultaneously over time leads to the updating innovations shown in part (b) of Proposition 1. The risk exposure innovation $s_t^\beta = \text{vec}\left(e_t u_t^\top \Sigma_u^{-1}\right)$ reflects that β can be understood as coefficients from regressing returns on factor innovations u_t while e_t is the corresponding error term. Therefore, e_t and u_t should be uncorrelated, which would mean that s_t^β is zero on average. If $e_t u_t^\top > 0$, we have $s_t^\beta > 0$, which leads to an increase in β (assuming the corresponding coefficient is positive). A positive influence function signals that there is some correlation between e_t and u_t, which means that e_t may possess some explanatory power for returns r_t that could also be associated with u_t. Since any (cross-sectionally) predictable variation in returns in a factor model should be due to the factor innovations, the MDAPM increases the beta to reduce the local correlation between the different innovations and to explain the additionally found predictable variation with a higher risk factor exposure.

The risk price updating scheme in the case of time-varying risk exposures is very similar to that with constant risk exposures. However, instead of updating risk prices solely due to projection errors $\left(\beta_{t-1}^\top \beta_{t-1}\right)^{-1} \beta_{t-1}^\top e_t$, these are scaled with $\left(1 - u_{t-1}^\top \Sigma_u^{-1} \lambda_t\right)$. If $u_t = 0$, the risk price updating behaves as in the case with time-constant exposures. However, if u_t and e_t are positive, the betas are increased to remove any possible correlation. In order not to increase the risk premium $\beta_t \lambda_t$ more than proportionally, the risk price movement is damped with the factor $\left(1 - u_t^\top \Sigma_u^{-1} \lambda_{t-1}\right)$.

2.4 Higher Order Moment Conditions

The set of conditional moment restrictions used in Sections 2.2 and 2.3 to derive the moment-based risk premium filter mainly represent the first-order moment conditions. Since the interpretation of factor volatility and idiosyncratic volatility is often connected towards risk, one could suspect that second order moment conditions should be informative for the risk premium updating. We therefore consider an extension of the MDAPM with time-varying volatilities that fits an extended set of conditional moment restrictions that is represented by the following conditional moment function:[3]

$$g_t(x_t;\theta_0) = \begin{pmatrix} u_t \\ \operatorname{vec}\left(u_t z_{t-1}^\top\right) \\ \operatorname{vec}\left(e_t u_t^\top\right) \\ e_t \\ \operatorname{vech}\left(u_t u_t^\top - \Sigma_u\right) \\ \operatorname{vech}\left(e_t e_t^\top - \Sigma_e\right) \end{pmatrix} \tag{22}$$

The following proposition presents the influence functions of the MDAPM based on the extended conditional moment conditions.

Proposition 2. Let x_t follow an MDAPM(p,q) based on the conditional moment function in (22). The influence function for $\vartheta_t = \left(\lambda_{t-1}^\top, \operatorname{vec}(\beta_{t-1})^\top, \operatorname{vech}(\Sigma_{u,t-1})^\top, \operatorname{vech}(\Sigma_{e,t-1})^\top\right)^\top$ is then given by

$$\begin{pmatrix} s_t^\lambda \\ s_t^\beta \\ s_t^{\Sigma_u} \\ s_t^{\Sigma_e} \end{pmatrix} = \begin{pmatrix} \left(\beta_{t-1}^\top \beta_{t-1}\right)^{-1} \beta_{t-1}^\top e_t \left(1 - u_t^\top \Sigma_{u,t-1}^{-1} \lambda_{t-1}\right) \\ \operatorname{vec}\left(e_t u_t^\top \Sigma_{u,t-1}^{-1}\right) \\ \operatorname{vech}\left(u_t u_t^\top - \Sigma_{u,t-1}\right) \\ \operatorname{vech}\left(e_t e_t^\top - \Sigma_{e,t-1}\right) \end{pmatrix} \tag{23}$$

[3] vech() refers to the half-vectorization of a symmetric matrix that stacks the columns of the lower triangular part into one column vector.

The general shape of MDAPM updating schemes for risk prices and exposures are not affected by additionally considering second order moment conditions. In particular, time-variation in idiosyncratic volatility, represented by $\Sigma_{e,t-1}$ does not affect risk premium dynamics. Thus, the derived MDAPM updating schemes are robust with respect to time-varying idiosyncratic volatility. However, factor volatility does impact risk premium dynamics as the time-varying factor covariance matrix $\Sigma_{u,t-1}$ enters the forcing innovations s_t^{λ} and s_t^{β}. The finding that $\Sigma_{u,t-1}$ does not enter the risk price update in the case of time-constant betas supports the view that volatility-based factor timing strategies (Barroso and Santa-Clara, 2015; Moreira and Muir, 2017) generate returns by predicting changes in factor exposures rather than factor risk prices.

Proposition 2 states that the influence functions for the covariance matrices $s_t^{\Sigma_u}$ and $s_t^{\Sigma_e}$ are given by the difference of squared errors and the current value of the corresponding covariance matrix. Moreover, this gives the corresponding updating schemes to follow VECH processes as in Bollerslev et al. (1988). This class of models is known for its large parameterization and the challenges it poses for estimation. That is particular troublesome in the MDAPM as the number of assets N is typically large. For practical reasons, it is therefore advisable to keep the idiosyncratic covariance matrix constant, as it does not affect the dynamics of risk premia. Since the number of factors K is typically much smaller than N, a VECH specification of factor innovations is likely to be feasible. Moreover, one could also consider restricted forms of VECH, such as the diagonal VECH or the BEKK specification of Engle and Kroner (1995).

2.5 Two-Step Estimation and Inference

The updating mechanism of time-varying parameters in the MDAPM was constructed under consideration of moment conditions. Therefore, it is natural to use the GMM method to estimate the static parameters as well. Since the number of such model parameters can be relatively large, I propose and use a two-step estimation approach.

2.5.1 First Estimation Step

We can decompose the conditional moment function (12) as

$$g_t(x_t;\theta_0) = \begin{pmatrix} g_{1t}(x_t;\phi_0,\phi) \\ g_{2t}(x_t;\theta_0) \end{pmatrix} \tag{24}$$

with

$$g_{1t}(x_t;\phi_0,\phi) = \begin{pmatrix} u_t \\ \text{vec}\left(u_t z_{t-1}^\top\right) \end{pmatrix} \qquad \text{and} \qquad g_{2t}(x_t;\theta_0) = \begin{pmatrix} \text{vec}\left(e_t u_t^\top\right) \\ e_t \end{pmatrix}. \tag{25}$$

Since the first set of moment conditions in g_{1t} only depends on the factor parameters ϕ_0 and ϕ, we can estimate those in a first step. Fortunately, the GMM estimator based on the moment conditions in g_{1t} is numerically equal to the ordinary least squares estimator for the parameters involved. We can therefore estimate the parameters of the factor dynamics with the following closed-form formula:

$$\left(\hat{\phi}_0, \hat{\phi}\right) = \left(\sum_{t=1}^{T-1} f_{t+1}\tilde{z}_t^\top\right)\left(\sum_{t=1}^{T-1} \tilde{z}_t\tilde{z}_t^\top\right)^{-1} \tag{26}$$

with $\tilde{z}_t^\top = (1, z_t^\top)$. Although the separate prior estimation of ϕ_0 and ϕ results in some loss of efficiency as these parameters also influence the moment conditions in g_{2t} via the factor innovation u_t, the two-stage approach also has two particular benefits. First, the possibility to derive closed formulas in the first stage reduces the computational effort, as these parameters would otherwise have to be computed numerically. Second, the prior fitting of factor innovations u_t seems to reduce overfitting the pricing errors $e_t = r_t - \beta_{t-1}\lambda_{t-1} - \beta_t u_t$. When estimating all parameters in the MDAPM simultaneously, the optimizer tends to introduce predictable variation in u_t in order to reduce e_t. This leads to a higher prediction error $r_t - \beta_{t-1}\lambda_{t-1}$, because the risk prices will not cover the conditional risk premium variation that is instead attributed to u_t.

2.5.2 Second Estimation Step

From the initial model parameters in θ_0, we still need to estimate the parameters of the updating scheme collected in $(\overline{\vartheta}^\top, \mathrm{vec}(A)^\top, \mathrm{vec}(B)^\top)$. With the introduction of time-varying risk premia, we have added additional static parameters. Since the static baseline model we started with was just identified, we now have to deal with the resulting underidentification. Therefore, we need to add additional moment constraints. As suggested by Creal et al. (2024), we define a vector of instruments $\zeta_t = (1, s_{t-1}^\top, (\vartheta_{t-1} - \overline{\vartheta})^\top)^\top \otimes I_{N(K+1)}$ and follow the approach of Hansen (1982) to minimize the GMM criterion given by

$$\min_{\overline{\vartheta}, A, B} \qquad \tilde{g}_T^\top \tilde{\Omega}_T \tilde{g}_T \tag{27}$$

with

$$\tilde{g}_T = \frac{1}{T}\sum_{t=1}^{T} \zeta_t g_{2t}(x_t; \vartheta_t, \tilde{\theta}_0) = \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} 1 \\ s_{t-1} \\ \vartheta_{t-1} - \overline{\vartheta} \end{pmatrix} \otimes g_{2t}(x_t; \vartheta_t, \tilde{\theta}_0) \tag{28}$$

where $\tilde{\Omega}_T$ is a positive definite matrix weighting the moment conditions. As usual, one could perform an initial estimation where the minimization problem (27) is solved with $\tilde{\Omega}_T = I$. These results can then be used to compute the long run variance of $g_t(x_t; \vartheta_t, \tilde{\theta}_0)$. The inverse of the long-run variance can be used as a weighting matrix in a second minimization of (27) to obtain more efficient estimates.

Using the lagged influences s_{t-1} and parameters $(\vartheta_{t-1} - \overline{\vartheta})$ as instruments reduces the first-order (cross-)autocorrelation of moments which are particularly important for inferring parameter updates. For example, if betas are constant, the additional moment conditions used would be $s_{t-1}^\lambda \otimes g_{2t}(x_t; \vartheta_t, \tilde{\theta}_0) = (\beta^\top \beta)\beta^\top e_{t-1} \otimes g_{2t}(x_t; \vartheta_t, \tilde{\theta}_0)$. Thus, the GMM estimator of the constant exposure MDAPM would adjust parameters to minimize the cross-correlation between lagged idiosyncratic innovations and contemporaneous moment conditions that include idiosyncratic and factor innovations as well as their products.

2.5.3 Inference

Creal et al. (2024) establish an asymptotic distribution theory for the general moment-based filtering framework. They show that under some high-level assumptions, it holds

$$\sqrt{T}\left(\hat{\theta}-\theta\right) \xrightarrow{d} \mathcal{N}\left(0, (G^\top \tilde{\Omega} G)^{-1} G^\top \tilde{\Omega} S \tilde{\Omega}^\top G (G^\top \tilde{\Omega} G)^{-1}\right), \tag{29}$$

where G is the limit of the gradient of (27) and the asymptotic covariance matrix $S = \sum_{j=-\infty}^{\infty} \mathbb{E}\left(g_t g_{t-j}^\top\right)$ of the moment conditions. The required high-level assumptions include that the filter of the time-varying parameter vector ϑ_t converges to a unique stationary and ergodic solution. For the case that betas are constant, i.e. $\vartheta_t = \lambda_t$, it can be shown straightforwardly that the stationarity assumption is fulfilled if $\|B - A\| < 1$ and $\|(\beta^\top \beta)^{-1}\| < \infty$.[4] Whereas the first inequality rules out explosive behavior of risk premia, the second one is met in absence of weak factors. In case of time-varying betas, such parameter restrictions are much more cumbersome to derive and out of the scope of this paper. The main challenge is that risk premia are then a product of different time-varying parameter vectors.

We use the result in (29) in the following application to compute standard errors for the parameter estimates, where the asymptotic covariance matrix is estimated with the heteroskedasticity and autocorrelation consistent estimator from Newey and West (1987) with Bartlett kernel as in Andrews (1991).

[4] $\|A\|$ denotes the spectral norm of a matrix A.

3 Simulation Study

In this section, the small sample performance of the MDAPM is evaluated with a Monte Carlo study on risk premium filtering.

3.1 Data-Generating Process

Assume there is one (cross-sectional) risk factor f_t whose factor mean can be predicted by the univariate series z_t with

$$f_t = z_{t-1} + u_t \tag{30}$$

where u_t follows a GARCH(1,1)-process with an unconditional variance of 18. The risk factor calibration is chosen to represent the features of the CRSP market return and the DGP can therefore be interpreted as a CAPM with dynamic coefficients. The factor mean predictor z_t is simulated from the AR(1) model given by

$$z_t = 0.5 + 0.98(z_{t-1} - 0.5) + \varepsilon_{z,t}, \qquad \varepsilon_{z,t} \sim \mathcal{N}\left(0, 0.137^2\right) \tag{31}$$

with its initial value z_1 drawn from the corresponding unconditional distribution. The process in (31) is calibrated in order to represent the moments and persistence of equity return predictors used in the following empirical application. Especially the highly persistent autoregressive parameter of 0.98 is often observed for factors considered as market return predictors (Campbell and Yogo, 2006).

The simulated returns r_t are derived from the beta representation according to (7) with N assets in the panel and T observations over time. Idiosyncratic innovations $e_{i,t}$ are drawn from a Student-t-distribution with $\nu = 6$ degrees of freedom and a variance of 4. The betas are simulated from a slow-moving processes given by

$$\beta_{i,t} = \overline{\beta}_i + 0.95(\beta_{i,t-1} - \overline{\beta}_i) + 0.05\varepsilon_{\beta_i,t}, \qquad \varepsilon_{\beta_i,t} \sim \mathcal{N}\left(0, 0.148^2\right). \tag{32}$$

where the N unconditional exposures $\overline{\beta}_i$ form an equidistant grid on the interval $[0.6, 1.5]$. Parameters in the process (32) are based on the results of the following empirical application. The unconditional beta interval is approximately the range observed for exposures when industry portfolio returns[5] are regressed on the market risk factor and a constant.

Since the DGP contains only one pricing factor, we need to simulate only one risk price λ_t. We consider four alternative settings for the dynamics of the risk price:

$$\text{Constant:} \qquad \lambda_t = 0.5 \tag{33}$$

$$\text{Cycle:} \qquad \lambda_t = 0.5 + \sin(2\pi t/T) \tag{34}$$

$$\text{Breaks:} \qquad \lambda_t = \begin{cases} 0.5, & \text{if } t \in [0, T/3] \\ -0.5, & \text{if } t \in (T/3, 2T/3] \\ 1.5, & \text{if } t \in (2T/3, T] \end{cases} \tag{35}$$

$$\text{AR:} \qquad \lambda_t = 0.5 + 0.98(\lambda_{t-1} - 0.5) + \varepsilon_{\lambda,t}, \qquad \varepsilon_{\lambda,t} \sim \mathcal{N}\left(0, 0.25^2\right) \tag{36}$$

The first constant risk price DGP serves as a benchmark to evaluate the performance of the dynamic MDAPM, since the actual process is static. Next, the cycle DGP reflects the idea that risk prices move with the business cycle. The third DGP with breaks is intended to mimic a situation in which unexpected news immediately change investors' perception of risk and, therefore, the price of risk demanded. Finally, the fourth process represents the situation where risk premia follow a highly persistent process, such as the betas in the employed DGP do.

[5] Although industry portfolios are not used in the empirical application, we use their exposure range for calibration in order to have a higher degree of dispersion in the betas to explore in simulations.

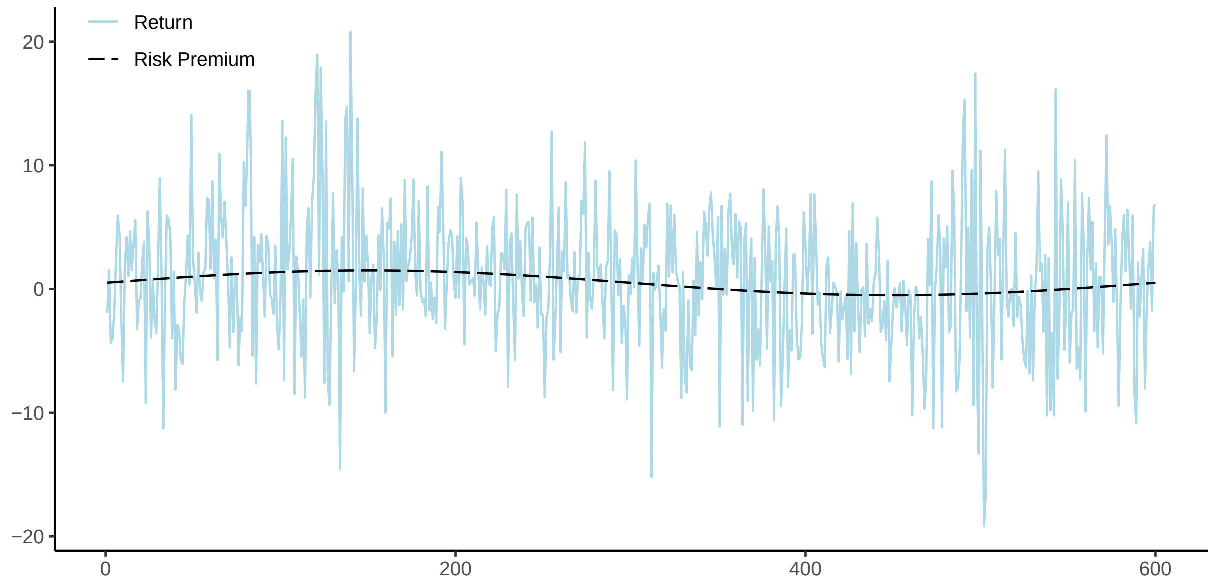

Figure 1: **Simulated Returns.** This figure shows simulated excess returns from an asset with unit (unconditional) risk factor exposure together with the cyclical risk premium λ_{t-1} from one draw of a panel with $N = 25$ assets and $T = 600$ time observations.

Figure 1 shows an example draw of the simulated excess return of an asset with unit (unconditional) risk factor exposure together with the cyclical risk premium λ_{t-1} from a panel of size $N = 25$ and $T = 600$. The cyclical variation in the conditional return expectation can be considered small compared to the variation in the realized return. This is consistent with the perception that stock returns have little predictability and poses a challenge to our method that attempts to filter out these almost diminishing dynamics. Also consistent with the stylized statistical facts of financial returns, the series shows volatility clustering, which in the DGP used comes from the GARCH residual of the pricing factor process in (30).

3.2 Simulation Results

In the following, we discuss the results of fitting a MDAPM with time-varying lambdas and betas to $S = 1000$ Monte Carlo replications of each of the four DGPs mentioned above.

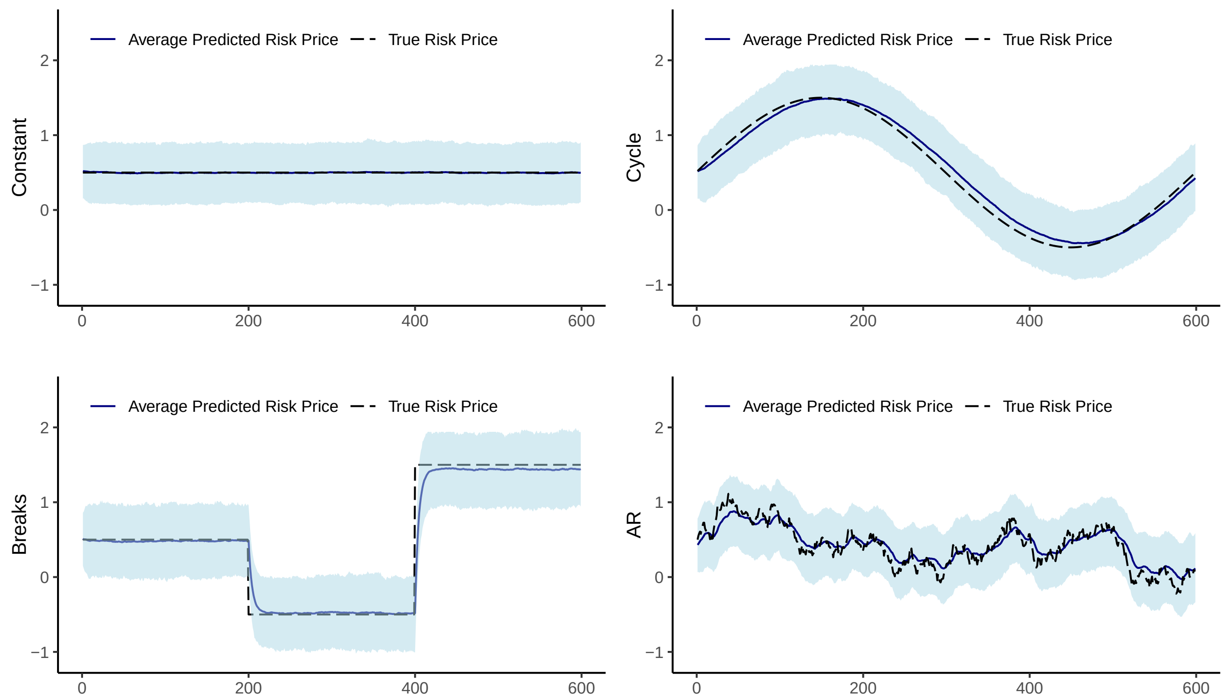

Figure 2: **Risk Price Predictions.** This figure shows, for a panel of size $N = 25$ and $T = 600$, the average risk price predicted by the MDAPM (solid line) together with the true risk price (dashed line). The shaded areas represent the 90 percent bands.

3.2.1 Predicted Risk Prices

Figure 2 shows, for a panel of size $N = 25$ and $T = 600$, the average risk price predicted by the MDAPM (solid line) together with the true risk price (dashed line). The shaded areas represent the 90 percent interquantile bands. We see that the MDAPM is able to adequately track the true risk price process in all four cases, on average with an interquantile band range of about 0.9. Given the rather low signal-to-noise ratio and the high variance of innovations, the filter uncertainty reflected by the bands can be considered rather low. In the case of a constant true risk price, the MDAPM is able to filter almost perfectly on average. In the cyclical case, the MDAPM tracks the true risk price with a short lag, as is typical for observation-driven models that process new information with a time lag. Similarly, the moment-based filter can react to breaks in the third DGP, but with a short delay. In the final AR case, the MDAPM faces the most difficulties. Although the filter correctly anticipates the direction of the true process and tracks it fairly well, it barely covers short-term peaks.

3.3 Pricing and Prediction Error Comparison

We further investigate the performance of the MDAPM with respect to pricing and prediction errors which are summarized in Table 1. The first metric for evaluating the pricing performance is the root mean squared pricing error (RMSE), which is computed as

$$RMSE_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \hat{e}_{i,t}^2}. \tag{37}$$

For comparability, we compute ΔRMSE as the difference between the RMSE of Fama and MacBeth (1973) regressions (hereafter FMB) assuming constant risk premia and the RMSE produced by the MDAPM. Thus, a positive ΔRMSE indicates that the MDAPM outperforms the FMB benchmark in terms of pricing errors.

Panel (a) in Table 1 shows the ΔRMSE averaged over portfolios and Monte Carlo replications for panels with different cross section sizes $N = 10, 25, 100$ and time series lengths $T = 300, 600, 1200$. We clearly see that the MDAPM outperforms the static FMB benchmark in every panel and DGP. This is particularly true for the constant risk price DGP, where the MDAPM benefits from the ability to account for time-varying betas and factor means compared to the FMB approach. As expected from Figure 2, the MDAPM performs worst for the fourth DGP with persistent autoregressive risk prices, although it still outperforms the benchmark. The latter benefits in this case especially from the low variability of risk prices. When comparing the performance of different panel sizes, we observe that the MDAPM performance improves monotonically with longer panels, i.e. higher T, but not necessarily with larger cross-sections, i.e. higher N. An explanation is that as the number of assets in the panel increases, additional parameters need to be estimated, namely the additional betas that are associated with the new assets. In contrast, increasing the length of the time series does not introduce additional parameters to the model and enables more

Table 1: Pricing and Prediction Error Comparison

The table shows difference root mean squared pricing errors (ΔRMSE) and root mean squared prediction errors (ΔRMSPE) of a MADPM with Fama and MacBeth (1973) regressions as benchmark averaged across assets and Monte Carlo replications. The nine simulated panels have different numbers of assets N, time observations T and are replicated 1000 times each.

	N=10			N=25			N=100		
	T= 300	600	1200	T=300	600	1200	T=300	600	1200
(a) Average ΔRMSE									
Const	0.0796	0.0948	0.1055	0.0925	0.1043	0.1096	0.0977	0.1035	0.1067
Cycle	0.1564	0.1913	0.2144	0.1636	0.1981	0.2072	0.1547	0.1562	0.1788
Breaks	0.2093	0.2322	0.2468	0.1988	0.2341	0.2525	0.1981	0.2016	0.2049
AR	0.0639	0.0939	0.1080	0.0808	0.1041	0.1117	0.0762	0.0934	0.1077
(b) Average ΔRMSPE									
Const	0.0031	0.0010	0.0004	0.0038	0.0014	0.0008	0.0044	0.0018	0.0010
Cycle	0.0499	0.0487	0.0488	0.0528	0.0493	0.0497	0.0554	0.0512	0.0502
Breaks	0.0634	0.0635	0.0629	0.0657	0.0627	0.0642	0.0662	0.0621	0.0637
AR	0.0084	0.0064	0.0046	0.0107	0.0077	0.0067	0.0111	0.0090	0.0071

accurate estimation due to improved data coverage.

The second metric we evaluate is the the root mean squared prediction error (RMSPE):

$$RMSPE_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_{i,t} - \hat{\beta}_{i,t-1}\hat{\lambda}_{t-1}\right)^2}. \tag{38}$$

This measure captures how well the model fits that the conditional expectation of r_{t+1} is given by $\beta_t\lambda_t$. Again, we construct a difference measure ΔRMSPE with an FMB baseline model to compare results from different panels. The averaged ΔRMSEs are shown in Panel (b) of Table 1. We observe that the MDAPM is also superior to the static FMB benchmark in every panel and DGP with respect to prediction errors. In contrast to the ΔRMSE results, we see that the performance of the MDAPM with respect to the static benchmark improves with increasing N. Hence, it seems that estimating the risk premium, whose dimension is unchanged, benefits from a richer cross-section, but other parameters are estimated with greater uncertainty, thus increasing idiosyncratic pricing errors. The rather counterintuitive result that the ΔRMSPEs decrease with T is technical. The beta processes in the DGP exhibit fairly high persistence, leading to large and long-lasting deviations from the unconditional mean. These deviations, which become more pronounced in longer samples, are generally more difficult to track than movements closer to the unconditional mean.

In summary, the MDAPM is capable of tracking the variation of risk premia with various dynamics in a setting with a realistically low signal-to-noise ratio. Its pricing performance, in particular, benefits from longer time series. Additionally, the risk premium prediction improves with larger cross sectional dimensions.

4 Empirical Application

The following empirical application examines the dynamics of risk premia that can be derived from the asset pricing moments in the 5-factor model of Fama and French (2015).

4.1 Data

Test assets are 32 equity portfolios sorted by size, operating profitability and investment. Stocks are first sorted into two groups according to size using NYSE median market cap breakpoints. Then each size group is allocated into four groups based on operating profitability that is measured as annual revenues minus cost of goods sold, interest expense, as well as selling, general, and administrative expenses divided by book equity. Finally, the eight groups are allocated each into four subgroups based on investment, measured as the change of total assets divided by toal assets of the previous fiscal year. The 32 monthly series are obtained from Kenneth French's online library and cover the period from January 1964 to June 2023. Thus, we work with a return panel with dimensions $N = 32$ and $T = 714$.

We consider the five factors from Fama and French (2015) to price the cross-section of

32 equity portfolios that are given by

$$f_t = (MKT_t,\ SMB_t,\ HML_t,\ RMW_t,\ CMA_t)^\top \tag{39}$$

where MKT is the excess return on the value-weighted equity market portfolio and the other factors are returns to long-short quantile portfolios with stocks sorted by market capitalization (SMB), book-to-market ratio (HML), profitability (RMW) and investment strategy (CMA).

As forecast instruments for the conditional factor mean we use the three-dimensional vector

$$z_t = (TSY10_t,\ TERM_t,\ DY_t)^\top \tag{40}$$

where TSY10 is the 10-year treasury yield and TERM is the term spread, calculated as the difference between the yields of the 10-year treasury note and the three-month treasury bill. Both series are obtained from the H.15 statistical release of the Board of Governors of the Federal Reserve System. The third forecasting factor DY is the dividend yield of the S&P 500 index. Evidence on equity return predictability from these factors can be found in Keim and Stambaugh (1986), Campbell (1987), Fama and French (1989), and Campbell and Thompson (2008) for long-run treasury yields and Campbell and Shiller (1988), Fama and French (1989), Campbell and Thompson (2008), and Cochrane (2008) for the term structure and dividend yields.

4.2 Empirical Model Specifications

The main specification to be considered is an MDAPM with time-constant betas. The parameter matrices A and B in the updating scheme are assumed to be diagonal in order to achieve an updating equation given by

$$\lambda_t^j = \overline{\lambda}^j + a_j^\lambda s_{j,t-1}^\lambda + b_j^\lambda \left(\lambda_{t-1}^j - \overline{\lambda}^j \right) \tag{41}$$

for every cross-sectional pricing factor $j = MKT, SMB, HML, RMW, CMA$. Note that the updating equations are parameterized along their long-run values $\overline{\lambda}^j$ which can be interpreted as risk prices of the corresponding factor j. The diagonalization mutes the impact of the influence function on the parameter updating of the other factors. However, excluding those effects in the present application does not crucially impair the model performance but rather yields a much more parsimonious model.

Another considered MDAPM specification allows for time-varying betas. We assume that the diagonal parameters for the beta update of the exposure of asset i to factor j are the same for each asset i. This assumption allows the model with beta dynamics to be estimated for larger panels and follows the assumption that exposures to the same factor follow similar dynamics. The resulting updating equations, parameterized along the long-run values $\overline{\beta}$, are given by

$$\text{vec}\left(\beta_{i,t}^j\right) = \text{vec}\left(\overline{\beta}_i^j\right) + a_j^\beta s_{i,j,t-1}^\beta + b_j^\beta \text{vec}\left(\beta_{i,t-1}^j - \overline{\beta}_i^j\right) \tag{42}$$

with parameters $\overline{\beta}^j$, a_j^β, and b_j^β, where $i = 1, \ldots, 10$ and $j = MKT, SMB, HML, RMW, CMA$.

Two established benchmark specifications are considered. The first benchmark is the unconditional risk price specification underlying classical Fama and MacBeth (1973) regressions. In line with Adrian et al. (2015), estimated innovations $\hat{u}_t$ from a VAR(1) model including the abovementioned factors are provided as pricing factors in order to account for the significant autocorrelation within the pricing factors. The second benchmark is a DAPM that explains risk price variations with the forecasting factors described above. This yields a regression equation given by

$$\lambda_t^j = \lambda_0 + \Lambda_1^{j,TSY10} TSY10_t + \Lambda_1^{j,TERM} TERM_t + \Lambda_1^{j,DY} DY_t \tag{43}$$

for each of the five cross-sectional risk factors $j = MKT, SMB, HML, RMW, CMA$. Estimation and inference for the DAPM is performed as described in Adrian et al. (2015), and I refer to them for more details.

4.3 Empirical Results

In the following, we present the empirical results of the application to the Fama-French 5-factor model regarding parameter estimates, filtered risk premia, and pricing errors.

4.3.1 Parameter Estimates

Parameter estimates for the risk premium updating schemes are shown in Table 2. The first row in the upper panel shows the estimated unconditional risk prices $\overline{\lambda}$ of the two MDAPM specifications. Risk prices of the benchmark model are presented in the first row of the lower panel. We see that the average price of market risk in the MDAPM differs only moderately from the one obtained in the static unconditional specification while some of the other unconditional risk prices differ more substantially without a clear direction. In particular, the CMA factor risk price more than halves relative to the unconditional benchmark and turns insignificant based on the standard errors shown in parentheses. Regarding the learning rates of the risk prices a^{λ}, we find that they are significantly positive for all five factors if betas are assumed to be constant. This means that their premia vary over time. If betas are allowed to vary over time, we find that the learning rates of RMW and CMA risk prices turn insignificant and the risk premium variation is captured by variation in risk exposures. Estimates of persistence rates b^{λ} indicate that risk prices vary with moderate persistence for all factors but HML. When risk exposures are time-varying, b^{λ} decreases slightly for all

Table 2: Risk Premium Parameter Estimates

This table shows estimates of risk premium parameter estimates for the five factors of Fama and French (2015) described in Section 4.1. The columns in the upper panel show results from a MDAPM with constant and time-varying betas. Time-varying betas are derived in a specification with unique learning and persistence rates per factor for the beta updating. The first five columns in the lower part provide estimates from a dynamic asset pricing model (DAPM) in line with Adrian et al. (2015). Forecasting factors are the 10-year treasury yield (TSY10), the term spread (TERM), and the dividend yield of the S&P 500 index (DY). The last five columns in the lower panel provide regression results of a constant (unconditional) risk price specification. MDAPM standard errors shown in parentheses are derived as described in Section 2.5. Errors for the DAPM estimates are adjusted for cross-asset correlation in the residuals and for estimation error of the time-series betas. GMM standard errors are reported for the unconditional specification. Test assets are 32 value-weighted equity portfolios sorted by size, operating profitability and investment with monthly returns denoted in percentages. The sample period is 1964:01 - 2023:06.

	MDAPM					MDAPM (t.-v. β)				
	MKT	SMB	HML	RMW	CMA	MKT	SMB	HML	RMW	CMA
$\overline{\lambda}$	0.612	0.230	0.424	0.252	0.100	0.527	0.171	0.314	0.360	0.249
	(0.156)	(0.110)	(0.158)	(0.103)	(0.088)	(0.164)	(0.109)	(0.210)	(0.137)	(0.135)
a^{λ}	0.501	0.255	0.092	0.098	0.070	0.115	0.120	0.129	0.090	0.055
	(0.050)	(0.073)	(0.027)	(0.040)	(0.028)	(0.036)	(0.043)	(0.032)	(0.061)	(0.048)
b^{λ}	0.896	0.907	0.515	0.845	0.927	0.819	0.845	0.772	0.835	0.796
	(0.033)	(0.045)	(0.226)	(0.092)	(0.033)	(0.092)	(0.098)	(0.093)	(0.152)	(0.182)
a^{β}						0.032	-0.015	0.025	-0.004	0.030
						(0.009)	(0.002)	(0.006)	(0.005)	(0.009)
b^{β}						0.941	0.983	0.858	0.943	0.902
						(0.031)	(0.004)	(0.058)	(0.122)	(0.042)

	DAPM					Unconditional				
	MKT	SMB	HML	RMW	CMA	MKT	SMB	HML	RMW	CMA
$\overline{\lambda}$	0.550	0.222	0.290	0.335	0.274	0.552	0.224	0.284	0.337	0.274
	(0.153)	(0.147)	(0.291)	(0.105)	(0.126)	(0.169)	(0.115)	(0.160)	(0.087)	(0.082)
TSY10	-0.243	-0.157	0.278	0.078	0.136					
	(0.084)	(0.058)	(0.088)	(0.044)	(0.041)					
TERM	0.203	0.055	0.205	0.133	-0.030					
	(0.135)	(0.093)	(0.131)	(0.070)	(0.066)					
DY	1.608	1.247	-1.529	-0.485	-0.771					
	(0.620)	(0.422)	(0.687)	(0.329)	(0.304)					

factors except HML, which actually becomes more persistent.

With respect to the learning and persistence rates for the beta series (a^{β} and b^{β}), we see that most betas except for RMW significantly vary over time. The persistence rates b^{β} are slightly higher than for risk prices with values around 0.9. Thus, the movement of risk exposures is somewhat slower than the movement of risk prices. This is reasonable since the test assets are characteristics-sorted portfolios whose exposures should change little or not at all.

The estimated coefficients of DAPM are comparable to those documented in Adrian et al. (2015), but the term spread appears to be less relevant in the considered 32 portfolios than it is in their combined equity and bond cross-section.

4.3.2 Filtered Risk Premia

Figure 3 shows the filtered risk prices of the MDAPM specification alongside the risk prices proposed by the DAPM and the unconditional model. For the MKT risk price in the upper panel of Figure 3, we see that the two dynamic approaches tend to move more or less in unison. In general, the DAPM risk price seems to be slightly higher in periods of high risk premia, such as the second half of the 1970s and the 2010s. Notably, the moment-based risk prices are more volatile. Moreover, we can also find that risk premia tend to increase in recessionary periods when markets are in a downturn. This countercyclicality has already been documented with regression-based approaches in Adrian et al. (2015), Gagliardini et al. (2016, 2020), and Chaieb et al. (2021). At the same time, we also see the pattern that market risk premia tend to fall at the beginning of crisis periods, as can best be seen around the global financial crisis. This pattern has also been documented in Jensen (2018), Gómez-Cram (2022) and Umlandt (2023). One reason could be that time-varying lambdas reflect not only the price of risk, but also the expected return. Thus, market risk prices may decline at the beginning of a recession due to lower expectations. In contrast to the results of the likelihood-based filter in Umlandt (2023), we see that although the market risk premium falls at the beginning of a recession, it rises quickly after the first drop, as can best be

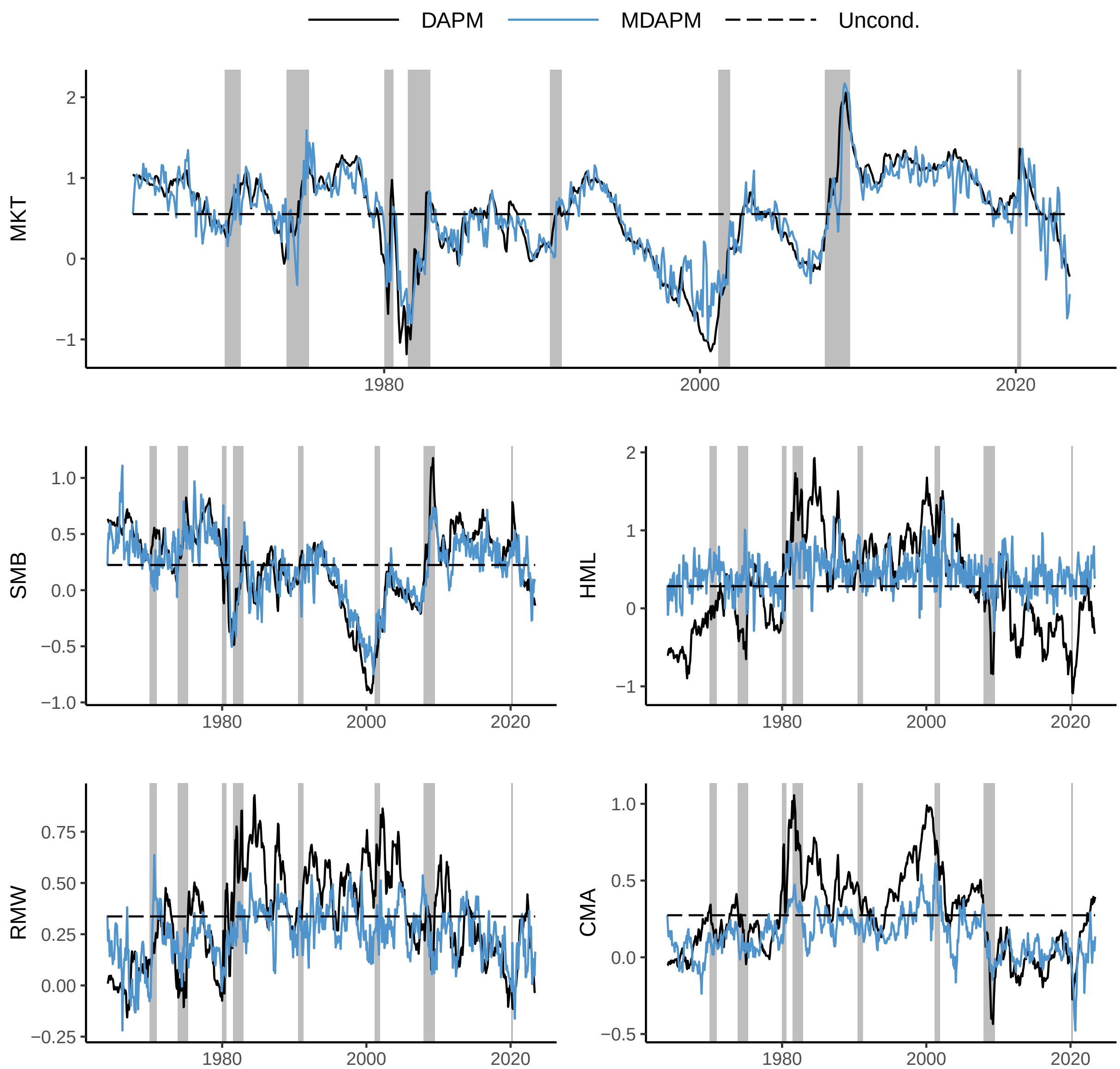

Figure 3: **Time-Varying Risk Prices.** This figure shows estimated risk prices for the five factors of Fama and French (2015) described in Section 4.1. It reports results from MDAPM with time-varying exposures and a DAPM from Adrian et al. (2015). Test assets are 25 value-weighted equity portfolios sorted on size and value with monthly returns denoted in percentages. The sample period is 1964:01 - 2023:06.

seen again in the global financial crisis. Thus, it appears that after an initial adjustment of expectations, adverse events cause investors to demand higher compensation for risk, in line with the previously found countercyclicality.

The lower panels of Figure 3 show predicted risk premia for the other four factors. While the SMB risk price dynamics implied by the MDAPM and DAPM are considerably similar, the two approaches suggest very different trajectories for the remaining three factors. The difference is particularly apparent for the HML risk price. While the DAPM suggests notably moving and persistent risk premia, those obtained from the MDAPM fluctuate rather volatile around an otherwise rather stable value that may have increased a bit in the middle of the sample. Consistent with the parameter estimates, both the MDAPM-implied risk premia for RMW and CMA show more persistent movement than that for HML. However, the movement is not as pronounced as in the regression-based results. Both factors seem to carry a rather stable risk premium from the 1980s until the global financial crisis. Especially the CMA premium seems to be close to zero after 2009.

One reason for the observation that the two approaches are mostly in agreement for market risk and the size risk premium, and only weakly so for the other factors, is that the risk price predictors used in the DAPM are specifically documented to predict market returns. These predictors appear to adequately span the information set for the conditional price of market risk, although they may not be sufficient to span the information set for the other pricing factors. Therefore, it is likely that more appropriate instruments can be found to study the risk premium dynamics of the non-market risk factors in a regression-based framework.

4.3.3 Pricing and Prediction Error Comparison

We evaluate the in-sample pricing and premium prediction performance according to two measures. The first measure for evaluating the pricing performance is the RMSE, as computed in (37). Columns 1 through 4 of Table 3 show the RMSE for the 32 portfolios individually as well as averaged over the entire set of test assets. From the averages, we can see that the moment-based MDAPM filter with constant exposures largely reduces the average RMSE from 1.924 in the unconditional model to 1.854, while the regression-based DAPM benchmark's RMSE is only reduced to 1.900. Thus, using the moment-based filter yields a roughly three times larger improvement than using the regression-based approach. Also from a portfolio-specific perspective, we find that the MDAPM crucially reduces the pricing error in comparison to DAPM and FMB for most portfolios. Note also that allowing risk exposures to vary over time, as seen in the results in column 2, yields additional small pricing error improvements compared to the constant β MDAPM specification. Therefore, the inclusion of time-varying risk prices seems to be more important than the inclusion of time-varying exposures. This is not too surprising given that the test assets are sorted portfolios.

The second metric we evaluate is the RMSPE defined in (38). Again, this measure captures how well the model fits that the conditional time $t-1$ expectation of r_t is given by $\beta_{t-1}\lambda_{t-1}$. This is because the conditional expectation of u_t should be zero. Because the variance of the factor innovation is typically relatively high compared to that of the idiosyncratic observations, the RMSPEs shown in columns 5 through 8 of Table 3 are considerably larger than the RMSEs. In addition, the difference between the conditional and the unconditional model specification is much smaller.

The smallest errors in predicting the risk premium, with an average of 5.531, are found for the MDAPM with constant β. This is considerably lower than the unconditional (5.649 on average) and regression-based dynamic (5.610 on average) benchmark. Again, note that while the absolute improvements appear small, the MDAPM produces a three times larger improvement over the unconditional benchmark than the DAPM. This is surprising as the DAPM estimates risk prices by regressing the earlier estimated product $\beta\lambda_t$ on the forecast factors z_{t-1}. Thus, the approach is particularly suited to adjusting risk prices in a way that minimizes the RMSPE. However, the constant exposure MDAPM seems to improve upon that while simultaneously reaching a better fit for other moments captured by the RMSE. The introduction of time-varying betas increases the RMSPE relative to the constant beta variant and falls between the RMSPE of the DAPM and the unconditional model. Thus, the slight improvement in RMSE from allowing time-varying betas seems to come at the cost of a deterioration in the risk premium prediction error.

Table 3: Root Mean Squared Pricing and Prediction Errors

The table shows root mean squared pricing errors (RMSE) and root mean squared prediciton errors (RMSPE) of different asset pricing model specifications. Pricing factors are the five Fama and French (2015) factors. MDAPM refers to results from specifications in which risk exposures are either constant or time-varying. DAPM refers to a dynamic asset pricing model specification according to Adrian et al. (2015) using the risk price predictors z_t defined in (40). Unc. refers to errors from constant lambda specifications estimated with Fama and MacBeth (1973) regressions. Test assets are 25 value-weighted equity portfolios sorted on size and value with monthly returns denoted in percentages. sivj refers to the portfolio of stocks in the intersection of the i-th quintile portfolio sorted on size and the j-th quintile portfolio sorted on value. The sample period is 1964:01 - 2023:06.

	RMSE				RMSPE			
	MDAPM		DAPM	Unc.	MDAPM		DAPM	Unc.
	β	β_t			β	β_t		
ME1.OP1.INV1	1.943	1.993	2.227	2.279	7.469	7.602	7.575	7.621
ME1.OP1.INV2	1.824	1.847	1.959	1.985	6.462	6.605	6.582	6.607
ME1.OP1.INV3	2.049	2.036	2.143	2.187	6.723	6.784	6.756	6.815
ME1.OP1.INV4	2.309	2.236	2.304	2.410	7.930	8.082	7.997	8.115
ME1.OP2.INV1	1.834	1.919	1.975	2.001	5.899	6.010	5.998	6.031
ME1.OP2.INV2	1.416	1.392	1.498	1.521	5.212	5.281	5.265	5.303
ME1.OP2.INV3	1.494	1.488	1.515	1.532	5.446	5.475	5.454	5.470
ME1.OP2.INV4	1.784	1.741	1.880	1.933	6.431	6.447	6.416	6.483
ME1.OP3.INV1	2.586	2.579	2.641	2.660	6.101	6.245	6.219	6.260
ME1.OP3.INV2	1.453	1.380	1.404	1.423	4.918	4.986	4.967	5.006
ME1.OP3.INV3	1.304	1.329	1.372	1.398	5.079	5.133	5.106	5.145
ME1.OP3.INV4	1.378	1.277	1.485	1.534	6.248	6.345	6.314	6.366
ME1.OP3.INV1	2.099	2.063	2.114	2.145	6.173	6.180	6.158	6.230
ME1.OP4.INV2	1.704	1.646	1.725	1.753	5.683	5.784	5.759	5.811
ME1.OP4.INV3	1.437	1.467	1.463	1.498	5.725	5.808	5.766	5.821
ME1.OP4.INV4	1.411	1.397	1.597	1.649	6.636	6.714	6.671	6.729
ME2.OP1.INV1	2.001	1.950	2.033	2.038	5.400	5.557	5.554	5.571
ME2.OP1.INV2	2.400	2.265	2.271	2.273	5.023	5.171	5.158	5.195
ME2.OP1.INV3	2.003	2.103	2.031	2.046	5.120	5.264	5.250	5.263
ME2.OP1.INV4	2.031	1.979	2.055	2.071	5.842	5.951	5.928	5.995
ME2.OP2.INV1	1.891	1.850	1.905	1.911	4.546	4.646	4.649	4.666
ME2.OP2.INV2	1.824	1.841	1.836	1.837	4.361	4.460	4.461	4.483
ME2.OP2.INV3	2.051	2.078	2.066	2.069	4.861	5.005	5.007	5.026
ME2.OP2.INV4	2.121	2.114	2.143	2.159	5.470	5.596	5.587	5.624
ME2.OP3.INV1	1.974	1.975	1.952	1.957	4.404	4.500	4.492	4.514
ME2.OP3.INV2	1.725	1.771	1.741	1.743	4.299	4.404	4.402	4.409
ME2.OP3.INV3	1.710	1.689	1.759	1.768	4.697	4.805	4.800	4.817
ME2.OP3.INV4	1.979	1.995	2.061	2.084	5.677	5.776	5.760	5.810
ME2.OP4.INV1	1.941	1.977	1.962	1.978	4.611	4.722	4.721	4.754
ME2.OP4.INV2	1.791	1.751	1.763	1.773	4.356	4.399	4.405	4.422
ME2.OP4.INV3	1.775	1.822	1.798	1.802	4.482	4.525	4.526	4.537
ME2.OP4.INV4	2.090	2.026	2.123	2.141	5.717	5.823	5.802	5.863
Avg	1.854	1.843	1.900	1.924	5.531	5.628	5.610	5.649

In summary, the error comparisons suggest that the moment-based approach can perform an estimation of dynamic asset pricing models with a crucially improved fit compared to both the conditional and the unconditional benchmark. It appears that the most substantial improvement comes from the presence of time-varying risk prices. Admittedly, it may be that better performance can be achieved by using individual updating schemes for the betas instead of specifying unique learning and persistence rates. However, this is challenging to implement as the number of parameters would increase massively and may tend to introduce more overfitting, as already suggested by the relatively high prediction errors of the specification used.

4.3.4 Alternative Factor Models and Cross-Sections

The Fama-French 5-factor model (FF5), together with the 32 test assets sorted by size (ME), operating profitability (OP) and investment (INV) is a suitable test application for the MDAPM because the factors cover a large portion of the cross-sectional variation and the portfolios have a strong factor exposure. We further investigate the potential of the MDAPM to uncover risk premia dynamics on a set of alternative factors and test assets. As alternative factor models we consider the CAPM that only includes the market factor MKT, the 3-factor model of Fama and French (1993) (FF3) that includes the factor MKT, SMB and HML, as well as the Carhart (1997) model (FFC) that adds a momentum factor to the three factors in FF3. The momentum signal is based on the prior returns in $t-12$ to $t-1$. Besides the already investigated 32 portfolio cross-section, we consider 32 portfolios sorted by size (ME), book-to-market ratio (BM) and operating profitability (OP), 32 portfolios sorted by size (ME), book-to-market ratio (BM) and investment (INV), as well as a merged cross section of 25 portfolios sorted on ME$\times$BM and 10 portfolios sorted by momentum (MOM).

Table 4 shows average RMSEs and RMSPEs for the alternative factor models and test assets. We find that for all panels and factor models except the CAPM, the MDAPM specifications perform best regarding the RMSE with the time-varying beta specification

Table 4: Error Comparisons for Alternative Factor Models and Cross-Sections

The table shows average root mean squared pricing errors (RMSE) and root mean squared prediciton errors (RMSPE) of different asset pricing model specifications and cross-sections. Factor models are the CAPM including only MKT, the FF3 model including MKT, SMB and HML, and the FFC model including the FF3 factor plus a momentum factor. MDAPM refers to results from specifications in which risk exposures are either constant or time-varying. DAPM refers to a dynamic asset pricing model specification according to Adrian et al. (2015) using the 10-year treasury yield (TSY10), the term spread (TERM), and the dividend yield of the S&P 500 index (DY) as risk price predictors. Uncond. refers to errors from constant lambda specifications estimated with Fama and MacBeth (1973) regressions. The sample period is 1964:01 - 2023:06.

	RMSE				RMSPE			
	MDAPM (const. β)	MDAPM (t.v. β)	DAPM	Uncond.	MDAPM (const. β)	MDAPM (t.v. β)	DAPM	Uncond.
(a) 32 ME×OP×INV								
CAPM	2.843	2.870	2.845	**2.831**	**5.570**	5.655	5.617	5.654
FF3	2.044	**1.985**	2.079	2.098	**5.538**	5.627	5.612	5.651
FFC	2.063	**1.992**	2.076	2.100	**5.533**	5.625	5.611	5.649
FF5	1.854	**1.843**	1.900	1.924	**5.531**	5.628	5.610	5.649
(b) 32 ME×BM×OP								
CAPM	3.415	3.439	3.422	**3.408**	**5.986**	6.075	6.041	6.077
FF3	2.458	**2.370**	2.484	2.504	**5.956**	6.053	6.035	6.074
FFC	**2.481**	2.537	2.484	2.509	**5.959**	6.041	6.034	6.074
FF5	2.348	**2.287**	2.365	2.389	**5.951**	6.059	6.033	6.073
(c) 32 ME×BM×INV								
CAPM	**2.976**	3.021	2.998	2.984	**5.591**	5.669	5.641	5.674
FF3	1.961	**1.926**	2.024	2.042	**5.566**	5.662	5.638	5.674
FFC	1.986	**1.974**	2.022	2.046	**5.564**	5.657	5.637	5.673
FF5	1.886	**1.854**	1.947	1.968	**5.562**	5.659	5.637	5.672
(d) 25 ME×BM + 10 MOM								
CAPM	**2.859**	2.904	2.881	2.867	**5.645**	5.744	5.711	5.752
FF3	1.921	**1.860**	1.978	1.997	**5.613**	5.725	5.709	5.751
FFC	1.717	**1.676**	1.770	1.799	**5.621**	5.726	5.706	5.749
FF5	**1.905**	1.951	1.942	1.967	**5.608**	5.723	5.706	5.749

producing slightly smaller pricing errors. With regard to the CAPM specification we see that the dynamic approaches are not able to improve upon the unconditional benchmark in (a) and (b) and decrease pricing errors only slightly in (c) and (d). This is not surprising given the rather low cross-sectional explanatory power of the CAPM with a cross-sectional R^2 of 17.25 percent (compared to a cross-sectional R^2 of 79.52 percent in FF5). The MDAPM uses cross-sectional regression errors to infer time dynamics, which are obscured by uncaptured cross-sectional variation in the CAPM and therefore provide a less clean signal of risk premium variation.

With respect to the RMSPE, we see that the MDAPM with constant β always performs best, even for the 1-factor CAPM. Thus, the uncaptured cross-sectional variation seems to be important for the MDAPM to improve the cross-sectional pricing, but not necessarily for the risk premium prediction. Consistent with the previous results, the performance of the time-varying beta specification falls between the two benchmarks.

Applying the FFC model to the 32 portfolio panels (a) to (c) yields a significantly positive risk premium for the momentum factor, but no improvement in pricing errors in the unconditional setting. Moreover, the inclusion of the additional factor in FFC does not noticeably affect the performance of the dynamic specifications compared to the FF3 model. In particular, in the MDAPM specifications, the risk price of the momentum factor seems to be quite stable.

Panel (d) shows pricing errors based on a test asset panel with a total of 35 portfolios, of which 25 are bivariate sorts with respect to ME and BM, and the other 10 are univariate sorts with respect to momentum. In contrast to the other panels, we see that the inclusion of the momentum factor in the FFC model clearly reduces pricing errors compared to the FF3 and FF5 models. However, with respect to RMSPE, we see that the risk premium prediction errors are not considerably affected by the inclusion of a momentum factor, as they are similarly high for FF3, FF5, and FFC. This means that the momentum factor explains portfolio returns cross-sectionally via its exposure to the term $\beta_t u_t$, as indicated by the lower RMSE. However, given the non-improved RMSPE, this exposure is not compensated by a higher risk premium that can be explained by the time variation in the risk prices of other factors. Thus, in line with the recent literature on factor momentum (Arnott et al., 2023; Ehsani and Linnainmaa, 2022), the MDAPM analysis reveals that the momentum premium can be explained by variation in the risk prices of other factors, but finds additional unpriced momentum exposure that explains why the inclusion of a momentum factor still improves cross-sectional pricing.

5 Conclusions

We introduced with the MDAPM a GMM-based dynamic asset pricing framework for linear factor pricing models. Time-varying risk premia are derived from an updating scheme that seeks a steepest descent improvement of the local GMM criterion function of a parsimonious set of asset pricing moments in the corresponding time period. It turns out that such a constructed updating mechanism adjusts risk prices according to regression errors from the cross-sectional regression performed in the second stage of the FMB procedure. In the case of time-varying betas, these are updated to enforce orthogonality of factor innovations and idiosyncratic innovations.

The MDAPM is applicable to a wide range of factor asset pricing models, does not require the specification of time series predictors, and does not require the specification of residual distributions. It is therefore a fairly robust alternative to the regression- and likelihood-based approaches of Adrian et al. (2015), Gagliardini et al. (2016), and Umlandt (2023), respectively. Estimation and inference can be performed in a standard GMM fashion.

Simulation results and an application to the Fama-French 5-factor model show that the MDAPM can substantially reduce pricing errors compared to static Fama and MacBeth (1973) regressions and the DAPM of Adrian et al. (2015). Filtered risk premia show a countercyclical pattern, with an initial decline at the beginning of crisis periods, and appear to differ crucially across factors.

Appendix

A Proofs

The following proofs of propositions use the following lemma. A proof of it can be found, for example, in Ronchetti and Trojani (2001) or Creal et al. (2024).

Lemma 1. *In case of the conditional moment condition* $\mathbb{E}_{t-1}\left[g_t(x_t;\vartheta_t,\tilde{\theta}_0)\right]=0$*, the influence function in* (16) *is given by*

$$s_t = -\left(\bar{G}_{\vartheta,t}^\top \Omega_{t-1}\bar{G}_{\vartheta,t}\right)^{-1}\bar{G}_{\vartheta,t}^\top\Omega_{t-1}g_t(x_t;\vartheta_t,\tilde{\theta}_0) \tag{A.1}$$

with

$$\bar{G}_{\vartheta,t} = \mathbb{E}_{t-1}^{\epsilon}\left(\frac{\partial g_t(x_t;\vartheta_t,\tilde{\theta}_0)}{\partial\vartheta_t^\top}\right)\Bigg|_{\epsilon=0} \tag{A.2}$$

where $\mathbb{E}_{t-1}^{\epsilon}$ *is the conditional expectation based on the measure* F_x^{ϵ}.

A.1 Proof of Proposition 1

(a) For $\vartheta_t=\lambda_{t-1}$ we derive

$$\bar{G}_{\vartheta,t} = \mathbb{E}_{t-1}^{\epsilon}\left(\frac{\partial g_t(x_t;\vartheta_t,\tilde{\theta}_0)}{\partial\vartheta_t^\top}\right)\Bigg|_{\epsilon=0} = \begin{pmatrix}0_{K\times K} & 0_{K\times NL} & 0_{K\times NK} & -\beta^\top\end{pmatrix}^\top \tag{A.3}$$

Using Lemma 1 and (A.3), we can derive the required influence function:

$$\begin{aligned} s_t = \frac{d\vartheta_t(F^{\epsilon})_x}{d\epsilon}\Bigg|_{\epsilon=0} &= -\left(\bar{G}_{\vartheta,t}^\top \Omega_{t-1}\bar{G}_{\vartheta,t}\right)^{-1}\bar{G}_{\vartheta,t}^\top\Omega_{t-1}g_t(x_t;\vartheta_t,\tilde{\theta}_0) && \text{(A.4)}\\ &= \left(\beta^\top\beta\beta^\top\beta\right)^{-1}\beta^\top\beta\beta^\top e_t && \text{(A.5)}\\ &= \left(\beta^\top\beta\right)^{-1}\beta^\top e_t && \text{(A.6)}\\ &= \left(\beta^\top\beta\right)^{-1}\beta^\top r_t-\lambda_{t-1}-u_t. && \text{(A.7)} \end{aligned}$$

(b) Given $\vartheta_t = \left(\lambda_{t-1}^\top, vec(\beta_{t-1})^\top\right)^\top$ we derive

$$\bar{G}_{\vartheta,t} = \begin{pmatrix} 0_{(NL+K)\times K} & 0_{(NL+K)\times NK} \\ 0_{NK\times K} & -\Sigma_u \otimes I_N \\ -\beta_t & -\lambda_t^\top \otimes I_N \end{pmatrix}. \tag{A.8}$$

We further find

$$\bar{G}_{\vartheta,t}^\top \Omega_{t-1} \bar{G}_{\vartheta,t} = \begin{pmatrix} 0_{K\times(NL+K)} & 0_{K\times NK} & -\beta_{t-1}^\top \\ 0_{NK\times(NL+K)} & -\Sigma_u \otimes I_N & -\lambda_t \otimes I_N \end{pmatrix} \begin{pmatrix} I_{M-N} & 0_{(M-N)\times N} \\ 0_{N\times(M-N)} & \beta_t\beta_t^\top \end{pmatrix}$$
$$\cdot \begin{pmatrix} 0_{(NL+K)\times K} & 0_{(NL+K)\times NK} \\ 0_{NK\times K} & -\Sigma_u \otimes I_N \\ -\beta_t & -\lambda_t^\top \otimes I_N \end{pmatrix} \tag{A.9}$$
$$= \begin{pmatrix} 0_{K\times(NL+K)} & 0_{K\times N} & -\beta_t^\top\beta_t\beta_t^\top \\ 0_{NK\times(NL+K)} & -\Sigma_u \otimes I_N & -\lambda_t \otimes \beta_t\beta_t^\top \end{pmatrix}$$
$$\cdot \begin{pmatrix} 0_{(NL+K)\times K} & 0_{(NL+K)\times NK} \\ 0_{NK\times K} & -\Sigma_u \otimes I_N \\ -\beta_t & -\lambda_t^\top \otimes I_N \end{pmatrix} \tag{A.10}$$
$$= \begin{pmatrix} \beta_t^\top\beta_t\beta_t^\top\beta_t & \lambda_t^\top \otimes \beta_t^\top\beta_t\beta_t^\top \\ \lambda_t \otimes \beta_t\beta_t^\top\beta_t & \Sigma_u^2 \otimes I_N + \lambda_t\lambda_t^\top \otimes \beta_t\beta_t^\top \end{pmatrix}. \tag{A.11}$$

Applying the block matrix Schur complement allows us to invert the matrix in (A.11):

$$\left(\bar{G}_{\vartheta,t}^\top \Omega_{t-1} \bar{G}_{\vartheta,t}\right)^{-1} = \begin{pmatrix} (\beta_t^\top\beta_t)^{-2} + \lambda_t^\top\Sigma_u^2\lambda_t(\beta_t^\top\beta_t)^{-1} & -\lambda_t^\top\Sigma_u^{-2} \otimes \beta_t(\beta_t^\top\beta_t)^{-1} \\ -\Sigma_u^{-2}\lambda_t \otimes (\beta_t^\top\beta_t)^{-1}\beta_t^\top & \Sigma_u^{-2} \otimes I_N \end{pmatrix}. \tag{A.12}$$

Moreover,

$$\bar{G}_{\vartheta,t}^{\top}\Omega_{t-1}g_t(x_t;\vartheta_t) = \begin{pmatrix} 0_{K\times(NL+K)} & 0_{K\times N} & -\beta_t^{\top}\beta_t\beta_t^{\top} \\ 0_{NK\times(NL+K)} & -\Sigma_u\otimes I_N & -\lambda_t\otimes\beta_t\beta_t^{\top} \end{pmatrix} \begin{pmatrix} u_t \\ \mathrm{vec}\left(u_t z_t^{\top}\right) \\ \mathrm{vec}\left(e_t u_t^{\top}\right) \\ e_t \end{pmatrix} \tag{A.13}$$

$$= \begin{pmatrix} -\beta_t^{\top}\beta_t\beta_t^{\top}e_t \\ -\mathrm{vec}\left(e_t u_t^{\top}\Sigma_u\right) - \lambda_t\otimes\beta_t\beta_t^{\top}e_t \end{pmatrix}. \tag{A.14}$$

Equations (A.11) and (A.14) together with Lemma 1 can be used to finally derive the influence function:

$$s_t = \begin{pmatrix} (\beta_t^{\top}\beta_t)^{-2} + \lambda_t^{\top}\Sigma_u^{2}\lambda_t(\beta_t^{\top}\beta_t)^{-1} & -\lambda_t^{\top}\Sigma_u^{-2}\otimes\beta_t(\beta_t^{\top}\beta_t)^{-1} \\ -\Sigma_u^{-2}\lambda_t\otimes(\beta_t^{\top}\beta_t)^{-1}\beta_t^{\top} & \Sigma_u^{-2}\otimes I_N \end{pmatrix}$$

$$\times \begin{pmatrix} -\beta_t^{\top}\beta_t\beta_t^{\top}e_t \\ -\mathrm{vec}\left(e_t u_t^{\top}\Sigma_u\right) - \lambda_t\otimes\beta_t\beta_t^{\top}e_t \end{pmatrix} \tag{A.15}$$

$$= \begin{pmatrix} \left(\beta_t^{\top}\beta_t\right)^{-1}\beta_t^{\top}e_t\left(1 - u_t^{\top}\Sigma_u^{-1}\lambda\right) \\ \mathrm{vec}\left(e_t u_t^{\top}\Sigma_u^{-1}\right) \end{pmatrix}. \tag{A.16}$$

□

A.2 Proof of Proposition 2

For $\vartheta_t = \left(\lambda_{t-1}^\top, \mathrm{vec}(\beta_{t-1})^\top, \mathrm{vech}(\Sigma_{u,t-1})^\top, \mathrm{vech}(\Sigma_{e,t-1})^\top\right)^\top$ we can derive

$$\bar{G}_{\vartheta,t} = \mathbb{E}_{t-1}^{\epsilon}\left(\frac{\partial g_t(x_t;\vartheta_t,\tilde{\theta}_0)}{\partial \vartheta_t^\top}\right)\Bigg|_{\epsilon=0} \tag{A.17}$$

$$= \begin{pmatrix} 0_{(NL+K)\times K} & 0_{(NL+K)\times NK} & 0_{(NL+K)\times K(K+1)/2} & 0_{(NL+K)\times N(N+1)/2} \\ 0_{NK\times K} & -\Sigma_u \otimes I_N & 0_{NK\times K(K+1)/2} & 0_{NK\times N(N+1)/2} \\ -\beta_t & -\lambda_t^\top \otimes I_N & 0_{N\times K(K+1)/2} & 0_{N\times N(N+1)/2} \\ 0_{K(K+1)/2\times K} & 0_{K(K+1)/2\times NK} & I_{K(K+1)/2} & 0_{K(K+1)/2\times N(N+1)/2} \\ 0_{N(N+1)/2\times K} & 0_{N(N+1)/2\times NK} & 0_{N(N+1)/2\times K(K+1)/2} & I_{N(N+1)/2} \end{pmatrix}. \tag{A.18}$$

Similar derivations as in the proof of Proposition 1 yield

$$\left(\bar{G}_{\vartheta,t}^\top \Omega_{t-1}\bar{G}_{\vartheta,t}\right)^{-1} = \begin{pmatrix} (\beta_t^\top\beta_t)^{-2} + \lambda_t^\top\Sigma_u^2\lambda_t(\beta_t^\top\beta_t)^{-1} & -\lambda_t^\top\Sigma_u^{-2}\otimes\beta_t(\beta_t^\top\beta_t)^{-1} & 0_K \\ -\Sigma_u^{-2}\lambda_t\otimes(\beta_t^\top\beta_t)^{-1}\beta_t^\top & \Sigma_u^{-2}\otimes I_N & 0_{N^2} \\ 0_{V\times K} & 0_{V\times N^2} & I_V \end{pmatrix} \tag{A.19}$$

with $V = N(N+1)/2 + K(K+1)/2$ and

$$\bar{G}_{\vartheta,t}^\top \Omega_{t-1} g_t(x_t;\vartheta_t) = \begin{pmatrix} -\beta_t^\top\beta_t\beta_t^\top e_t \\ -\mathrm{vec}\left(e_t u_t^\top \Sigma_u\right) - \lambda_t \otimes \beta_t\beta_t^\top e_t \\ \mathrm{vech}\left(u_t u_t^\top - \Sigma_{u,t-1}\right) \\ \mathrm{vech}\left(e_t e_t^\top - \Sigma_{e,t-1}\right) \end{pmatrix}. \tag{A.20}$$

Equations (A.19) and (A.20) together with Lemma 1 prove the proposition.

□

References

Adrian, T., Crump, R.K., Moench, E., 2015. Regression-based Estimation of Dynamic Asset Pricing Models. Journal of Financial Economics 118, 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, 1931–1973.

Andrews, D.W.K., 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. Econometrica 59, 817–858.

Ang, A., Kristensen, D., 2012. Testing Conditional Factor Models. Journal of Financial Economics 106, 132–156.

Arnott, R.D., Kalesnik, V., Linnainmaa, J.T., 2023. Factor Momentum. The Review of Financial Studies 36, 3034–3070.

Barroso, P., Santa-Clara, P., 2015. Momentum has its Moments. Journal of Financial Economics 116, 111–120.

Bollerslev, T., Engle, R.F., Wooldridge, J.M., 1988. A Capital Asset Pricing Model with Time-Varying Covariances. Journal of Political Economy 96, 116–131.

Campbell, J.Y., 1987. Stock Returns and the Term Structure. Journal of Financial Economics 18, 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, 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, 1509–1531.

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

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

Chaieb, I., Langlois, H., Scaillet, O., 2021. Factors and Risk Premia in Individual International Stock Returns. Journal of Financial Economics 141, 669–692.

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

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

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

Creal, D., Koopman, S.J., Lucas, A., Zamojski, M., 2024. Observation-Driven Filtering of Time-Varying Parameters using Moment Conditions. Journal of Econometrics 238, 105635.

Ehsani, S., Linnainmaa, J.T., 2022. Factor Momentum and the Momentum Factor. Journal of Finance 77, 1877–1919.

Engle, R.F., Kroner, K.F., 1995. Multivariate Simultaneous Generalized Arch. Econometric Theory 11, 122–150.

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

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

Fama, E.F., French, K.R., 2015. A Five-Factor Asset Pricing Model. Journal of Financial Economics 116, 1–22.

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

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

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

Gagliardini, P., Ossola, E., Scaillet, O., 2020. Estimation of Large Dimensional Conditional Factor Models in Finance, Handbook of Econometrics, Volume 7A, pp. 219–282.

Giroux, T., Royer, J., Zerbib, O.D., 2024. Empirical Asset Pricing with Score-Driven Conditional Betas. Journal of Financial Econometrics , nbae007.

Gómez-Cram, R., 2022. Late to Recessions: Stocks and the Business Cycle. The Journal of Finance 77, 923–966.

Haddad, V., Kozak, S., Santosh, S., 2020. Factor Timing. Review of Financial Studies 33, 1980–2018.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 2011. Robust Statistics: The Approach based on Influence Functions. John Wiley & Sons.

Hansen, L.P., 1982. Large Sample Properties of Generalized Method of Moments Estimators. Econometrica 50, 1029–1054.

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

Jensen, C.S., 2018. Essays on Asset Pricing. Doctoral Thesis. Frederiksberg: Copenhagen Business School (CBS).

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

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are Covariances: A Unified Model of Risk and Return. Journal of Financial Economics 134, 501–524.

Lettau, M., Pelger, M., 2020. Estimating Latent Asset-Pricing Factors. Journal of Econometrics 218, 1–31.

Lustig, H., Verdelhan, A., 2012. Business Cycle Variation in the Risk-Return Trade-Off. Journal of Monetary Economics 59, 35–49.

Moreira, A., Muir, T., 2017. Volatility-Managed Portfolios. Journal of Finance 72, 1611–1643.

Neuhierl, A., Randl, O., Reschenhofer, C., Zechner, J., 2023. Timing the Factor Zoo. URL: https://papers.ssrn.com/abstract=4376898, doi:10.2139/ssrn.4376898.

Newey, W.K., West, K.D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55, 703–708.

Ronchetti, E., Trojani, F., 2001. Robust Inference with GMM Estimators. Journal of Econometrics 101, 37–69.

Umlandt, D., 2023. Score-Driven Asset Pricing: Predicting Time-Varying Risk Premia based on Cross-Sectional Model Performance. Journal of Econometrics 237, 105470.