Erasmus University Rotterdam
Erasmus School of Economics
Master Thesis Quantitative finance

Good and Bad Betas in Bull and Bear Markets: A Regime-Dependent Intertemporal CAPM

Ajay Lachmansingh (619062)

Supervisor:	dr. Erik Kole
Second assessor:	dr. Dick van Dijk
Date final version:	6th March 2024

Abstract

We evaluate an extended version, of the Intertemporal Capital Asset Pricing Model (ICAPM) based on a Markov switching VAR (MSVAR) model, ICAPM-MS, to assess its effectiveness in explaining the variation in returns of size and value sorted portfolios. The ICAPM-MS, introduced by Bianchi (2020), allows for the estimation of discount rates and cash-flow news terms through an MSVAR model. The constructed MSVAR model identifies recession (as defined by the NBER) and stable states of the US market. Consequently, the decomposed market shocks are estimated, assuming that long-term investors make an inference of the current and future states of the market. As demonstrated by Bianchi (2020), the ICAPM based on an specific MSVAR framework that identifies the Great Depression, yields greater explanatory power compared to the traditional ICAPM. Generalizing the MSVAR such that it identifies more common events like recessions, could offer insights on more broader characteristics of long-term investor behavior. Based on the results of various asset pricing tests using a rolling sample window, the ICAPM-MS can be more effective in explaining the variation in the cross-section of size and value sorted portfolios, however for specific periods and sample sizes. The improvements in performance with respect to a traditional ICAPM occurred during the transition of a recession regime to a stable regime. These results indicate that the assumption that the investor understands the changing dynamics of the market during such a transition is of importance. While the performance of the ICAPM-MS improved during these periods with respect to the ICAPM, the difference in performance measures is however not significantly large, and the overall performance of the ICAPM-MS is not high. The performance of the ICAPM-MS is also dependent on the window, suggesting that the constructed market factors do not sufficiently encapsulate systemic risk over an extended period.

Contents

Chapter 1

Introduction

Asset pricing theory is based on the assumption that, in an efficient market, the returns on an investment in a specific asset act as compensation for the systemic risk involved in investing in the asset. Based on this assumption, if all relevant risk factors of the asset are understood and known, they can be related to expected future returns. This approach has proven to be successful as financial institutions utilize asset pricing models to make investment decisions. Said models are of interest as they provide a quantitative framework for evaluating investment opportunities and constructing optimal portfolios. Describing risk factors and constructing asset pricing models that explain the cross-sectional variation in asset returns remains vital, given the dynamic nature of financial markets. Additionally, describing in what context, and under which specifications the models have a high performance is essential, such that the applicability and limitations of these models are well understood.

Recent research by Bianchi (2020) explores a particular asset pricing model, the two beta Intertemporal Capital Asset Pricing Model (ICAPM) introduced by Campbell and Vuolteenaho (2004), an extension of the CAPM that incorporates the expectation of long-term investors about future investment opportunities. The ICAPM is an extension of the traditional CAPM, as it assumes that investors take into account the future prospects of the assets they hold. To the contrary, the CAPM assumes that investors have constant expectations of future investment opportunities. An important finding by Bianchi (2020) is that the ICAPM, which additionally assumes that investors anticipate crises like the Great Depression, performs significantly better.

My aim is to evaluate the two beta ICAPM, where investors consider regime switching, and to understand under what circumstances and model specifications such a model performs well in explaining sectional variations in asset returns. More specifically, the question I wish to answer reads:

How does the explanatory power of the two beta ICAPM that considers recessions and stable *regimes on size and value sorted portfolios, compare to the two beta ICAPM and the CAPM that do not consider different regimes at all?.*

This question is addressed across various model specifications for different samples. This paper contributes to the literature by offering insights into systematic and period-specific investor sentiment within the ICAPM framework. This is of importance as such a model may explain investment behaviour in general, as opposed to ad-hoc factor models that mainly relate returns to specific firm characteristics. Additionally, the relevance of ICAPM is emphasized by Nagel (2013), as he concludes that variations of the ICAPM are able to explain the abnormalities in asset returns well for modern samples. This conclusion is drawn based on the comparison of a wide range of common factor models.

The state variables used to construct the MSVAR model are excess market returns, the term yield spread, the small value spread, the dividend-price ratio, and the price-earnings ratio. I compare this ICAPM to a benchmark ICAPM based on a regular VAR model that does not include inferences on regimes. Additionally, a CAPM model serves as a benchmark.

The performance is evaluated based the ability of the models to explain the variation in the cross-section of size-and value-sorted portfolio returns. Following the Fama-Macbeth two-pass approach, I construct test statistics including the GLS R^2 and p-values of a GLS test that examines whether the pricing errors are close to zero. The analysis assesses the models' capability to capture period-specific investor sentiment by constructing factor models using a 35-year rolling window. The capability to capture systemic investor sentiment is evaluated based on the same procedure on a full (modern) sample ranging from 03-1963 until 12-2022.

Findings suggest improvements of the MSVAR-based ICAPM during transitions from recession regimes to stable regimes, based on the GLS R^2 and the p-values of the pricing error test, which is not rejected during such periods at a 90% confidence level for this model. However, the overall performance of the model and the results of the full sample regression indicate a relatively low overall performance.

Chapter 2

Literature

2.1 The two factor Intertemporal CAPM

Empirical evidence has shown that the CAPM fails to explain the cross-sectional variation in asset returns since the 1960s. Fama and French (1992) show that the presence of anomalies like the size- and value effect contribute to the limited performance since they cannot be explained by the CAPM beta. This indicated that asset returns are exposed to other sources of risk. As a consequence, many other risk factors have been introduced that contribute in explaining the dispersion in average asset or portfolio returns.

One of the other fallacies of the CAPM is the inability to capture the long-term expectations of investors about future returns. The Intertemporal CAPM (ICAPM) by Merton (1973) accounts for this. In his paper, he derives an intertemporal model based on the behaviour of long-term investors who maximize their utility and account for uncertainty in future investment opportunities. It is concluded that the inherent assumption by the CAPM, that investors have a two-period horizon or assume constant investment opportunities over time, is unrealistic. Campbell and Shiller (1988) further describe how innovations in market returns are due to changes in expectations for future cash-flows as well as future discount rates. A decomposition method of market innovations into cash-flow and discount rate news terms is developed by Campbell (1991). Building upon the idea, that shocks in market returns is actually the aggregate of shocks in expectations of future discount rates and cash-flows, Campbell and Vuolteenaho (2004) argue that the CAPM beta should be split into a cash-flow beta and a discount rate beta, resulting in the two beta ICAPM. The rationale behind this is that investors recognize that news about future cash-flows and news about the future discount rate have different implications for future investment opportunities. While news about a decrease in future cash-flows and an increase in future discount rates both result in a contraction of the market, the discount rate news is associated with improved investment opportunities in the future. Cash-flow shocks are regarded as permanent shocks to wealth. Under the assumption that the discount rate news correlates with improved investment opportunities in the future, its corresponding risk premium should be lower. Hence, the negative cash-flow news is considered 'bad' and negative discount rate news 'good', or rather, less bad. The importance of this decomposition of risk sources is demonstrated by Campbell and Vuolteenaho (2004) as they find that the two risk factors are weakly correlated. Further results show differences in discount rate and cash-flow betas for different portfolios. They find that the two beta ICAPM is more successful than the CAPM in explaining the cross-sectional spread in average returns of value-size sorted portfolios.

2.2 Identifying regimes using the MSVAR model

The aforementioned decomposition method by Campbell (1991) assumes that investors predict market returns through a VAR(1) model consisting of excess market returns and macroeconomic variables. Based on the estimated VAR model, at period t, predictions are made of the market returns at $t+1$ as well as the long-run expectation of market returns. At period $t+1$, the market return shock can be defined as the difference in predicted and actual market returns. The long-run expectation of the market returns is re-estimated such that changes in this expectation can be related to innovations in discount rate expectations. Since both market shocks and discount rate shocks are now estimated, the cash-flow shock can be backed out as well.

This described method relies on a static VAR model, imposing that the investor expects constant economic conditions, disregarding changing dynamics due to factors such as business cycle effects and structural changes. As is pointed out by Campbell and Vuolteenaho (2004), the assumption of constant volatility of market returns may be an issue given empirical evidence on potential drawbacks in forecasting ability due to the presence volatility clustering. Engle (1982) and Bollerslev (1986) show the importance of capturing the heteroskedasticity in the conditional variance of returns via (G)ARCH models. Other methods assume the existence of different regimes, e.g. bull and bear markets or recession and expansion regimes. An example is the Threshold AutoRegressive (TAR) model, that allows for time-varying parameters within an autoregressive model that are dependent on the current regime (Tong, 1978). Similar is the Markov Switching VAR (MSVAR) model, introduced by Hamilton (1989). A general Markov Switching model assumes that regime-switching is governed by a hidden Markov process. The regime, or state, is latent. Therefore, at each period t, probabilities that each of the regimes prevails, are estimated. By estimating regime-specific model parameters along with probabilities for each regime, a mixture model can be constructed. While these regime-specific models may be linear, the eventual mixture model may be nonlinear depending on the estimated parameters of the model. In an MSVAR model, these regime-specific linear models are VAR models.

The terms bull and bear markets imply that there are two distinct states of the market. Empirical evidence supports this specification of market regimes (Ang & Timmermann, 2012). The two regimes are often characterized by high/low mean asset returns in combination with low/high volatility. Allowing for more regimes is done by Guidolin and Timmermann (2007) and Guidolin and Ono (2006), such that more specific states, e.g crashes and recovery states, can be recognized. To identify bull and bear markets, it is useful to allow for a time-varying mean, autoregressive coefficients and volatility within an MSVAR model (Ang & Timmermann, 2012). Results by Kole and Van Dijk (2017) show that including additional macro-financial variables to a Markov Switching model that already contains a volatility proxy as indicator does not have much of an effect on the regime probabilities. This implies that the volatility in returns dominates the market regime identifying process. In the same manner, Ang and Timmermann (2012) find that regimes are primarily identified by variations in volatility.

2.3 An ICAPM based on an MSVAR model

The previously discussed VAR decomposition method is generalised by Bianchi (2020) such that discount rate and cash-flow news terms are constructed through MSVAR models instead of regular VAR models. This extension allows for investors to take into account possible regimes changes as they predict future market returns. The MSVAR model estimated by Bianchi (2020) identifies the Great Depression regime. Based on the estimated smoothed probabilities, this regime does not reoccur for a long period but prevails again during the financial crisis for a brief moment. Results show that the two beta ICAPM, that considers the existence of this regime, have increased explanatory power. This conclusion is drawn based on a cross-sectional regression on size-value sorted portfolios using a 35-year rolling window. The R^2 of this regression is systematically and significantly higher than the R^2 of the regular ICAPM. These results suggest that the perception of investors of risk and future investment opportunities is dynamic and relies on their inference of the current and future state of the economy. It is concluded that, as long as the Great Depression or the Great Recession are included in the sample, the ICAPM is able to explain the cross section of asset returns. The performance of the ICAPM drops significantly whenever neither of the crises are considered. It is therefore interesting to assess if such results can also be found for an ICAPM that is based on an MSVAR that identifies recessions and stable regimes in general. In support of this hypothesis are the results by Gordon and St-Amour (2000). A consumption-based asset pricing model (C-CAPM) is developed such that risk aversion of agents is regime dependent. They find that their estimated preference regimes move in parallel to the bull and bear markets to a certain degree. Similar conclusions are drawn by Maio (2013), where a conditional ICAPM is constructed by splitting the cash-flow beta into a constant and a time-varying component. The time-varying beta is scaled by variables such as the CPI, making its fluctuations correlate with the business cycle. It is found that the conditional ICAPM is able to explain the spread in momentum sorted portfolios well, and that the scaled cash-flow beta plays a significant role to achieve this result. Asymmetry in behaviour of shocks in returns is shown by Kole and van Dijk (2023), but from a different perspective. They formulate an MSVAR model as an extended linear non-Gaussian VAR model considering bull and bear markets as the possible regimes. They find evidence that the shocks have a regime dependent effect on the Markov-switching process. It is concluded that the effects of the shocks vary in strength depending on the current state and the direction of the shock.

The decomposition method by Campbell (1991) relies on the investors' expectations of the future market. This expectation is modeled based on the level variables (intercepts and AR-coefficients) of the estimated VAR models. As is discussed before, the volatility of the returns often dominate the regime switching process, i.e. regimes are identified mainly based on differences in volatility. A consequence is that the ICAPM based on the MSVAR may closely resemble the traditional ICAPM if the regime-specific level variables are similar. This is accounted for by Bianchi (2020) by assuming different level regimes as well as variance (or volatility) regimes. The level regimes and variance regimes are further assumed to be independent from each other. The investors therefore make projections of the market while considering these level regimes.

In this paper, I develop an MSVAR based ICAPM, following the methodology described by Bianchi (2020). The aim is to assess the regime-dependent nature of the ICAPM, given the discussed evidence in favor of this specification. Various modifications to the ICAPM models constructed by Campbell and Vuolteenaho (2004) and Bianchi (2020) are considered, such that recession regimes are identified. Bianchi (2020) assumes independent level- and variance regimes. An issue with this method is that investors do not consider the variance regimes to identify bull and bear markets. This may be an issue as bear markets, or recessions in general are more dynamic and less persistent than the Great Depression regime. A compromise is made by allowing for four regimes, with each regime being a combination of a set of two level parameters and two variance parameters. Given this specification, the level regimes and variance regimes are dependent, allowing for the variance to affect the course of the level regimes. An issue with this method is that investors do not consider the variance regimes to identify bull and bear market.

Chapter 3

Data

The variables of the ICAPM estimated via an MSVAR model, the ICAPM-MS, are the excess market returns, the term yield spread, the small value spread, the dividend price ratio (D/P ratio) and the price earnings ratio (P/E ratio). Due to the design of the ICAPM-MS, the state variables serve as indicators of the business cycle as well predictors of the excess market returns. Both properties coincide to large degree, as the state variables capture shifts in investment behavior. Investors may adjust their portfolios based on expectations of economic conditions, influencing market returns. The term yield spread, defined as the difference between the interest rates of short- and long term government bonds, is a common business cycle predictor. The term yield spread is typically positive during times of economic stability and the spread is mainly caused by differences in duration risk. During recessions however, various factors such as monetary policies and changes in risk perspectives cause elevated short term interest rates and may even result in an inverted yield spread. The P/E ratio is the ratio between the level and earnings of companies in the S&P 500 index. This variable is a predictor of market returns, as an increase in the price of the market, given constant growth, is associated with lower long-term expected returns (Shiller, 1996). Predictive power of the D/P, which is the ratio of the annual paid out dividends and the level of the S&P 500 ratio of the market, is shown by Campbell and Shiller (1988) and Fama and French (1988). More recent evidence by Park (2010) shows that there is some predictive power as long as the D/P ratio is stationary, which depends on the sample period. The small value spread is the difference in log book-to-market ratios of small value and small growth stocks. The predictive ability of this variable arises due to the presence of the value anomaly. Small value portfolios are associated with higher expected returns than small growth portfolios. An increased small value spread implies better relative performance of small growth portfolios, which in turn implies lower future expected market returns.

Portfolios sorted on size and value are available in the library of Kenneth. R. French (French, 2012). The dataset provides the monthly average value weighted returns of the sorted portfolios. The excess markets returns are available in the library as well. Portfolio returns ranging from 03-1963 until 12-2022 are used for the asset pricing tests. The excess market returns are computed as the difference between the log return on the CRSP value-weighted stock index and the rate on 1-month T-bills. The term yield spread is calculated using data available on Global Financial Data, as the difference between 10-year constant-maturity yield and 3-Month T-bills, in percentage points. The small value spread is constructed by the log of the ratio of book-to-market/market equity (BE/ME) ratio of small value portfolios and small growth portfolios. These portfolios are formed annually in June based on sorts on size and book equity-to-market ratio. Firms considered small have a market equity lower than the median NYSE market equity in June. These small firms are further categorized based on their book equity-to-market equity ratio. Specifically, small growth firms are identified as the 30% of stocks with the lowest BE/ME ratio, while small value firms are those falling within the 30% of firms with the highest BE/ME ratio. The D/P ratio is defined as the log of the ratio of the annual paid out dividends and the level of the S&P 500. The annual dividends at month t is therefore constructed as the sum of the monthly dividends at the sum of the dividends of month $t-11$ until month t. Furthermore, the P/E ratio is defined as the log of the ratio between the level of the S&P 500 index and earnings of companies in the S&P 500 index. The data of the state variables ranges from 01-1928 until 12-2022. Table 1 shows summary statistics of the state variables. Statistics are determined using monthly dating ranging from 01-1928 until 12-2022. The presence of skewness along with large excess kurtoses indicate dynamics that deviate from normality. The strong autocorrelations highlight the importance of the autoregressive terms in the models.

3.1 stationarity

The standard VAR model, which is used as a benchmark, assumes stationary variables. Therefore, variables deemed nonstationary are differenced in both the VAR and MSVAR models, as the VAR model relies on the state variables being stationary. While the MSVAR model does not necessarily rely on individual variables to be stationary, this way the models consist of the same specification of the state variables. The Augmented Dicky Fuller (ADF) test is used to test for the presence of a unit root in each individual variables. If a variables is deemed to be nonstationary, it is differenced by the moving average of the six previous months.

The MSVAR model does not rely on the assumption of stationary individual variables. However, mean-square stability of the MSVAR process or a time series process in general is important to ensure that the model is robust and the assumptions on the statistical properties hold. There is mean-square stability if both the first- and second moment of the process are asymptotically finite. The test for mean-square stability is described in the methodology.

Table 1: Summary statistics variables

	Mean	Std.	Skewness	Kurtosis	Autocorr.
r_M^e	0.65	5.38	0.16	11.37	0.10
TY	1.44	1.10	-0.18	2.61	0.95
VS	1.73	0.34	1.26	4.26	0.98
DP	-3.42	0.48	-0.59	2.41	0.99
PE	2.85	0.41	0.73	2.27	0.99

Note: This table shows the mean and standard deviation in percentage points as well as the skewness, kurtosis, and the autocorrelation of the excess market returns (r_M^e), term yield spread (TY), small values spread (VS), the D/P ratio (DP), and the P/E ratio (PE).

Chapter 4

Methodology

4.1 The MSVAR model

A multivariate VAR(1) model consisting of n variables that assumes m potential regimes can be described by

$$\mathbf{y}_t = \mathbf{c}_{S_t} + \Phi_{S_t}\mathbf{y}_{t-1} + \Lambda_{S_t}\boldsymbol{\epsilon}_t, \tag{1}$$

with $\epsilon_t \sim N(0, I_n)$, where $\mathbf{y}_t$ is a vector of n variables, $\mathbf{c}$ a vector of length n consisting of intercepts, and $\mathbf{\Phi}$ is an $n \times n$ matrix that contains the autoregressive coefficients. The error terms are multiplied with the $n \times n$ matrix $\Lambda_{S_t}\epsilon_t$ to introduce heteroskedasticity to the model, with $\Lambda_{S_t}\Lambda'_{S_t}$ equal to the covariance matrix Σ_{S_t}. The terms that contain the subscript S_t vary per regime $S_t = 1, \ldots, M$.

Under the assumption that the regime process S_t evolves according to a time-invariant first order Markov chain, the probability of regime j prevailing exclusively depends on the current regime i. Thus the probability of entering regime j on time t conditional on the all past information until time $t-1$, I_{t-1} only depends on the prevalent regime at time $t-1$ and is given by

$$P_{ij} = P(S_t = j|I_{t-1}) = P(S_t = j|S_{t-1} = i) = p_{ij}. \tag{2}$$

The transition matrix P contains all transition probabilities.

Assuming two regimes as an example, we denote $\boldsymbol{\xi}_{t|t} = (1, 0)'$ if the true state at time t is 1 and $\boldsymbol{\xi}_{t|t} = (0, 1)'$ if the true state at time t is 0, given all information available time t. Due to the states being latent, the true states are not known. Therefore, an inference is made regarding the likelihood that the states prevail at time t, conditional on all available information up until time t, which is denoted by $\hat{\boldsymbol{\xi}}_{t|t}$. Element i of $\hat{\boldsymbol{\xi}}_{t|t}$, $\hat{\boldsymbol{\xi}}_{i,t|t} = P(S_t = i|I_t)$ for $i = 0, 1$.

4.1.1 The EM algorithm

The EM algorithm by Dempster, Laird and Rubin (1977) is an iterative algorithm that alternates between an expectation step, that updates the inferences made of the regimes prevailing, and a maximisation step, that updates the parameters of the regime-specific models. The expectation step consists of a prediction step, a filter step by Hamilton (2010) and the Kim (1994) smoother step. At the prediction step, the state at time $t+1$ is estimated, conditional on all available information up until time t. The transition probabilities are used and $\hat{\boldsymbol{\xi}}_{t+1|t}$ can be calculated by

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \mathrm{P}\hat{\boldsymbol{\xi}}_{t|t}. \tag{3}$$

At the filter step, $\hat{\boldsymbol{\xi}}_{t+1|t}$ is updated by incorporating the additional information of the next observable $\mathbf{y}_{t+1}$. Given two states, the updated estimation of the state $\hat{\boldsymbol{\xi}}_{t|t-1}$ is defined as

$$\hat{\boldsymbol{\xi}}_{t|t} = \begin{bmatrix} P(St = 1|I_{t-1}, \mathbf{y}_t) \\ P(St = 0|I_{t-1}, \mathbf{y}_t) \end{bmatrix} \tag{4}$$

and is calculated by

$$\hat{\boldsymbol{\xi}}_{t|t} = \frac{\mathbf{f}(\mathbf{y_t}) \odot \hat{\boldsymbol{\xi}}_{t|t-1}}{\boldsymbol{\iota}_2' \left[\mathbf{f}(\mathbf{y_t}) \odot \hat{\boldsymbol{\xi}}_{t|t-1}\right]'}, \tag{5}$$

with $\mathbf{f}(\mathbf{y_t})$ a vector of regime-specific pdf's and $\boldsymbol{\iota}_2$ a vector of ones with length 2. This can naturally be extended to multiple regimes. The following expressions and derivations are not constrained to two regimes. By performing predicting and filtering steps alternately, the predicted and filtered state can be obtained for the full timeframe of the dataset.

What follows is the smoothing step. The smoothed estimate $\hat{\boldsymbol{\xi}}_{t|T}$ corresponds to the expectation $\mathrm{E}[\xi_t|I_T]$. Using the tower property, this can be written as

$$\hat{\boldsymbol{\xi}}_{t|T} = \mathrm{E}[\mathrm{E}[\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, I_T]|I_T]. \tag{6}$$

The inner expectation can be expressed as

$$\mathrm{E}[\boldsymbol{\xi}_t|\boldsymbol{\xi}_{t+1}, I_T] = \hat{\boldsymbol{\xi}}_{t|t} \odot \mathrm{P}' \left[\boldsymbol{\xi}_{t+1} \oslash \hat{\boldsymbol{\xi}}_{t+1|t}\right] \tag{7}$$

and this expression can be used in equation 6 such that we get

$$\hat{\boldsymbol{\xi}}_{t|T} = \mathrm{E}\left[\hat{\boldsymbol{\xi}}_{t|t} \odot \mathrm{P}' \left[\boldsymbol{\xi}_{t+1} \oslash \hat{\boldsymbol{\xi}}_{t+1|t}\right] |I_T\right]. \tag{8}$$

By recognizing that the expectation of $\boldsymbol{\xi}_{t+1}$ given I_T is $\hat{\boldsymbol{\xi}}_{t+1|T}$, we rewrite equation 8 as

$$\hat{\boldsymbol{\xi}}_{t|T} = \hat{\boldsymbol{\xi}}_{t|t} \odot \mathrm{P}' \left[\hat{\boldsymbol{\xi}}_{t+1|T} \oslash \hat{\boldsymbol{\xi}}_{t+1|t}\right]. \tag{9}$$

Estimates of $\hat{\boldsymbol{\xi}}_{t+1|T}$ can be found by first performing all prediction and filter steps until time T. At the last filter step, an estimate of $\hat{\boldsymbol{\xi}}_{T|T}$ is calculated. Subsequently, using equation 9 the estimate $\hat{\boldsymbol{\xi}}_{T-1|T}$ can be derived since all terms on the right hand side have been calculated. This recursive step can be iterated to obtain smoothed estimates of the states for each period.

We estimate the optimal parameters $\hat{\theta}$ that maximises the log-likelihood $\log(f(\mathbf{y}|\boldsymbol{\theta}))$. The optimal parameters can be derived via

$$\hat{\theta} = \tilde{\mathrm{E}}\left[_{\boldsymbol{\theta}} \log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta})\right]. \tag{10}$$

The joint probability density function $f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta})$ cannot simply be written as the product of the individual density functions $\prod_{t=1}^{T} f(\mathbf{y}_t|\boldsymbol{\theta})$, due to the serial correlation between consecutive observations of $\mathbf{y}$. By applying Bayes' rule, the joint pdf can however be written as the product of conditional pdf's. Using this approach, the joint pdf can be expressed as follows:

$$f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta}) = \left[\prod_{t=2}^{T} f(\mathbf{y}_t|s_t; \boldsymbol{\theta}) P(S_t = s_t|s_{t-1}; \boldsymbol{\theta})\right] f(\mathbf{y}_1, s_1; \boldsymbol{\theta}). \tag{11}$$

The expectation of the log-likelihood function given by equation 10 is maximized for a given set of parameters, $\hat{\theta}_{\mathbf{0}}$, i.e.

$$\tilde{\mathrm{E}}[\log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta})|\boldsymbol{\theta}_{\mathbf{0}}] \tag{12}$$

is maximized and subsequently the expectation of the log-likelihood function using the estimated parameters is taken for the next iteration. The likelihood function can be reformulated as

$$f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta}, \mathrm{P}, \boldsymbol{\rho}) = \prod_{t=1}^{T} \left[\prod_{i,j=1}^{n} (f_i(\mathbf{y}_t) p_{ij})^{\delta_{ij}(t)}\right] \left[\prod_{j=1}^{n} \rho_j^{\delta_j(0)}\right], \tag{13}$$

such that the log-likelihood becomes

$$\log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta}, \mathrm{P}, \boldsymbol{\rho}) = \sum_{t=1}^{T}\sum_{i,j=1}^{n} \delta_{ij}(t)\log(f_i(\mathbf{y}_t)p_{ij}) + \sum_{j=1}^{n}\delta_j(0)\log\rho_j. \quad (14)$$

Here, $\delta_{ij}(t)$ is a random variable that is 1 if $S_t = j, S_{t-1} = i$. Similarly, $\delta_j(0)$ is 1 if $S_0 = j$. $\boldsymbol{\rho}$ corresponds to the regime probabilities at t is 0, i.e. $\boldsymbol{\rho} = \boldsymbol{\xi}_0$. Taking the expectation of equation 14 yields

$$\tilde{\mathrm{E}}[\log f(..)] = \sum_{t=1}^{T}\sum_{i,j=1}^{n} p^*_{ij}(t)\log(f_i(\mathbf{y}_t)p_{ij}) + \sum_{j=1}^{n} p^*_j(0)\log\rho_j. \quad (15)$$

Here $p^*_j(0) = P(S_t = j|I_T)$, is simply the smoothed estimate of the regime probabilities at $t = 0$. This expression of the expectation of the log-likelihood requires estimations of transition probabilities $p^*_{ij}(t) = P(S_t = j, S_{t-1} = i|I_T)$, which are not defined by the previous steps. The smoothed step therefore includes an additional step that reads

$$\mathrm{P}^*(t) = P\dot{(}\hat{\xi}_{t|T}\hat{\xi}'_{t-1|t-1}) \oslash (\hat{\xi}'_{t|t-1}\boldsymbol{\iota}'_n), \quad (16)$$

with the matrix $\mathrm{P}^*(t)$ containing each element $p^*_{ij}(t)$. The maximisation step $k+1$ consists of finding the parameters $\hat{\boldsymbol{\theta}}_{k+1}, \mathrm{P}_{k+1}, \boldsymbol{\rho}_{k+1}$ that maximise $\tilde{\mathrm{E}}[\log f(\mathbf{y}_{1:T}, \boldsymbol{\xi}_{1:T}|\boldsymbol{\theta}_k, \mathrm{P}_{\mathbf{k}}, \boldsymbol{\rho}_k)]$.

The maximization problem for the parameters is solved by setting the partial derivative of equation 15, w.r.t each respective parameter, to zero. By introducing the Lagrange multiplier κ and adding a term that enforces that $\sum_{j=1}^{n}\rho_j = 1$ we find that the optimal parameter for $\boldsymbol{\rho}$ is simply the smoothed estimate of the regime probabilities. For P_{k+1} we also include the Lagrange multiplier that enforces that $\sum_{i=1}^{n} p_{il} = 1$. Solving the partial derivative per regime m results in

$$p_{ml,k+1} = \frac{\sum_{t=1}^{T} p^*_{ml}(t)}{\sum_{t=1}^{T} p^*_l(t-1)}. \quad (17)$$

Given that $\sum_{i=1}^{n} p^*_{ij} = p^*_j$, to find $\boldsymbol{\theta}_{k+1}$ we solve

$$0 = \frac{d\sum_{t=1}^{T}\sum_{m=1}^{M} p^*_m(t)\log f(\mathbf{y}_t|\boldsymbol{\theta}_{m,k+1})}{d\boldsymbol{\theta}_{m,k+1}}. \quad (18)$$

We assume normality such that for the distribution of the parameters of regime m holds:

$$\log f(\mathbf{y_t}|\boldsymbol{\theta}_m) \propto \frac{1}{2}K(\log(2\pi) + \log|\Sigma_m| + \mathbf{u}_{mt}(\mathbf{c}, \boldsymbol{\Phi})'\Sigma_m^{-1}\mathbf{u}_{mt}(\mathbf{c}, \boldsymbol{\Phi})). \quad (19)$$

The elements of $\mathrm{P}^*(t)$ do not have a subscript $k+1$ for ease of notation. These elements, along with the other estimates of the expectation step, can be assigned the subscript $k+1$, as they occur at EM step $k+1$. By further recognizing that $p^*_m(t) = \hat{\xi}_{m,t|T}$, the numerator in equation 18 is formulated as

$$\frac{1}{2}\sum_{t=1}^{T}\sum_{m=1}^{M}\hat{\boldsymbol{\xi}}_{m,t|T}\left[K(\log(2\pi)+\log|\Sigma_m|+\mathbf{u}_{mt}(\mathbf{c},\boldsymbol{\Phi})'\Sigma_m^{-1}\mathbf{u}_{mt}(\mathbf{c},\boldsymbol{\Phi}))\right], \tag{20}$$

with $m = 1,\dots,M$ regimes, $\hat{\boldsymbol{\xi}}_{mt|T}$ the smoothed probability of regime m, Σ_m the covariance matrix of regime m and $\mathbf{u}_{mt}(c,\Phi)$ equal to $\mathbf{y}_t - \mathbf{c}_{S_t} + \Phi_{S_t}\mathbf{y}_{t-1}$, using the notation of the VAR model specified in equation 1. This specification assumes the variance parameters Σ_m and level parameters $\mathbf{u}_m$ are constrained to be jointly dependent on the same regime m.

4.1.2 The updating steps under the assumption of a combination of level- and variance regimes

To construct an MSVAR model that allow $m\dot{n}$ regimes that are combinations of m level- and n variance parameters, the VAR model given by equation 1 can be modified to

$$\mathbf{y}_t = \mathbf{c}_{S_t^m} + \Phi_{S_t^m}\mathbf{y}_{t-1} + \Lambda_{S_t^n}\epsilon_t, \tag{21}$$

where we both assume level regimes $S_t^m = 1,\dots,M$ and volatility regimes $S_t^n = 1,\dots,N$. This model assumes $m\dot{n}$ states, with $\hat{\boldsymbol{\xi}}_{mn,t|t}$ denoting the smoothed joint probability $P(S_t^m, = m, S_t^n = n|I_{t-1}, y_t)$.

The transition matrix P is initialised as the Kronecker product of the transition matrices of the individual Markov chains P^m and P^n. The level- and variance regimes are not assumed to follow separate Markov chains, therefore during optimization P is not constrained to be the Kronecker product of P^m and P^n anymore. Equation 20 is generalized to

$$-\frac{1}{2}\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{n=1}^{N}\hat{\boldsymbol{\xi}}_{mn,t|T}\left[K(\log(2\pi)+\log|\Sigma_n|+\mathbf{u}_{mt}(\mathbf{c},\boldsymbol{\Phi})'\Sigma_n^{-1}\mathbf{u}_{mt}(\mathbf{c},\boldsymbol{\Phi})\right]. \tag{22}$$

Here, $\mathbf{u}_{mt}(\mathbf{c},\boldsymbol{\Phi})$ is defined as $\mathbf{y}_t - (\mathbf{c}_{S_t^m} + \Phi_{S_t^m}\mathbf{y}_{t-1})$. The maximization step entails solving the systems of partial derivatives w.r.t Σ and $\boldsymbol{\gamma}$. Where $\boldsymbol{\gamma}$ consists of a single vector of all level coefficients, i.e. $\boldsymbol{\gamma} = (\mathbf{c}', \boldsymbol{\Phi})'$ following Krolzig (2013). This formulation allows for solving the partial derivatives in a computationally convenient way. This can be achieved by redefining the MSVAR model as

$$\mathbf{y}_t = \sum_{m=1}^{M}\xi_{mt}X_{mt}\boldsymbol{\gamma} + \Lambda_{S_t^n}\epsilon_t, \tag{23}$$

with X_{mt} a $K \times R$ matrix given K state variables and R the total amount of level coefficient which is equal to M times $(K+1)$. X_{mt} structures the regressors via

$$X_{mt} = [(\boldsymbol{\iota}_m, \tilde{\mathbf{y}}_{t-1,m}) \otimes I_k], \tag{24}$$

with $\boldsymbol{\iota}_m$ a $1 \times M$ vector of zeros and with the m'th element equal to 1, and $\tilde{\mathbf{y}}_{t-1,m}$ a $1 \times MK$ vector defined as $\boldsymbol{\iota}_m \otimes \mathbf{y}'_{t-1}$.

Eventually, for maximazation step $k+1$, the partial derivative with respect to the structural parameters γ_{k+1} can be rewritten as

$$\boldsymbol{\gamma}_{k+1} = (\boldsymbol{X}'W^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'W^{-1}(1_M \otimes \mathbf{y}). \tag{25}$$

X is a $MTK \times R$ matrix constructed by stacking all matrices X_{mt}. W^{-1} is a block-diagonal matrix with each diagonal element given by $\sum_{n=1}^{N} \hat{\Xi}_{mn} \otimes \Sigma_{n,k}^{-1}$. $\mathbf{y}$ is defined as $(\mathbf{y}_1{}')', \ldots, (\mathbf{y}_T{}')'$ This is equivalent to the solution of a GLS regression. In addition to weights assigned based on the inverse of the covariance matrices Σ, it can be observed from W^{-1} that each regression is weighted with their corresponding estimated smoothed probability. Solving the partial derivative w.r.t the covariance matrix leads to

$$\hat{\Sigma}_n = \sum_{m}^{M-1} \frac{1}{\hat{T}_{mn}} \mathbf{u}_m(\boldsymbol{\gamma_k})' \hat{\Xi}_{mn} \mathbf{u}_m \boldsymbol{\gamma_k}, \tag{26}$$

with $\hat{T}_{mn}$ as the sum of $\boldsymbol{\xi}_{mn,t|T}$ of all t, $\hat{\Xi}_{mn}$ a diagonal matrix with element t equal to $\boldsymbol{\xi}_{mn,t|T}$, and $\mathbf{u}_m\boldsymbol{\gamma} = (\mathbf{u}_{m0}, ...\mathbf{u}_{mT})$. Notice that there is no simultaneous update of all shape parameters γ and Σ. Both parameters use an expression of the other at the maximization step. As can be observed by the subscripts k at the right hand side of equations 25 and 26, the estimated parameter of Σ at step k is used to update γ at step $k{+}1$ and vice versa. Simultaneously updating the parameters is not possible, making the expression in equation 18 slightly inconsistent with the actual process.

For the algorithm, all parameters are initialized with a degree of randomness, ensuring they are not excessively unreasonable as well as satisfy the constraints implied by the model. Below is the pseudocode shown that calculates the optimal parameters.

Algorithm 1 MSVAR

the dataset $\mathcal{T}$ runs from 1928-01 until 2022-12

Set $l^* = -10^8$
Set $threshold = 10^-2$
for $n = 1, 2, \ldots$ until $Nruns$ **do**
  Initialize parameters P_0, $\boldsymbol{\lambda_0}$ and $\boldsymbol{\theta_0}$
  Set $l_0 = -10^8$
  for $i = 1, 2, \ldots$ until $Niterations$ **do**
    for $t \in \mathcal{T}$ **do** store in $\mathcal{T}$
      Run Hamilton filter/smoother to obtain estimates of state probabilities $\boldsymbol{\lambda_i}$
    end for
    for $t \in \mathcal{T}$ **do** store in $\mathcal{T}$
      Calculate parameters P_i, $\boldsymbol{\theta_i}$ using $\boldsymbol{\lambda_i}$ and $\boldsymbol{\theta_{i-1}}$
    end for
    Calculate log-likelihood l_i
    if $l_i - l_i - 1 < threshold$
    $l_i = l$
    P_i, $\boldsymbol{\lambda_i}$ and $\boldsymbol{\theta_i} = P$, $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$
    Break
  end for
  if $l > l^*$ $l = l^*$
  P, $\boldsymbol{\lambda}$ and $\boldsymbol{\theta} = P^*$, $\boldsymbol{\lambda}^*$ and $\boldsymbol{\theta}^*$
  Return $P^*, \boldsymbol{\lambda}^*, \boldsymbol{\theta}^*$

4.1.3 Mean-square stability

As is stated bofore, there is mean-square stability if both the first- and second moment of the process are asymptotically finite. R. E. Farmer, Waggoner and Zha (2009) show that even with unstable, but inpersistent, regimes the MSVAR process can be stable. Under the assumption of ergodicity and covariance stationarity of the innovation terms, the MSVAR model is mean-square stable if and only if is second order stationary in the limit (Costa & Dufour, 2005). Covariance stationarity can be proven by showing that the spectral radius of the matrix

$$\Upsilon(P \otimes I_{n^2}) < 1, \tag{27}$$

where P is the transition matrix and I_{n^2} a $n^2 \times n^2$ identity matrix. Υ is the block diagonal bdiag$[\Phi_1, \ldots, \Phi_{mn}]$ of the autoregressive coefficients (Bianchi, 2016), (Kole & van Dijk, 2023).

4.1.4 Standard errors of the estimates

As the EM-algorithm maximizes the likelihood function, standard assumptions on the maximum likelihood estimators hold. Therefore, if the regularity conditions discussed by Krolzig (2013) are satisfied, the estimated parameter vector $\hat{\boldsymbol{\theta}}$ converges in distribution to a normal distribution, i.e.

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}_0^{-1}), \tag{28}$$

where $\boldsymbol{\theta}_0$ is the true parameter vector. The most suitable estimation of the information matrix $\mathcal{I}_0$ in the case of many parameters is the outer product of the gradient of the conditional likelihood

$$\hat{\mathcal{I}}_G(\hat{\boldsymbol{\theta}}) \equiv \frac{1}{T} \sum_{t=1}^{T} \mathbf{g}_t(\hat{\boldsymbol{\theta}}) \mathbf{g}_t(\hat{\boldsymbol{\theta}})', \tag{29}$$

where the conditional score $\mathbf{g}_t(\hat{\boldsymbol{\theta}})$ can be expressed as

$$\mathbf{g}_t(\hat{\boldsymbol{\theta}}) \equiv \frac{\partial \ell(\mathbf{y}_t | I_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial \ell(I_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - \frac{\partial \ell(I_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \tag{30}$$

The conditional score is decomposed in this equation by Hamilton (1993) to avoid solving a long sequence of derivatives. To see this, Kole (2019) shows that, for given shape parameter vector(e.g the set of VAR parameters) $\boldsymbol{\lambda}$,

$$\frac{\partial \ell(I_t; \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \sum_{\tau}^{t} \frac{\partial \log \mathbf{f}'_\tau}{\partial \boldsymbol{\lambda}} \boldsymbol{\xi}_{\tau|t}. \tag{31}$$

The distribution function $\boldsymbol{f}_t = P(\mathbf{y}_t | s_t, \boldsymbol{\lambda})$ is in this case the normal distribution, also denoted in equation 19. Using the recursive relation in equation 30, Kole (2019) expresses the score as

$$\frac{\partial \ell(\mathbf{y}_t | I_{t-1}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial \log \mathbf{f}'_t}{\partial \boldsymbol{\lambda}} (\boldsymbol{\xi}_{t|t}) + \sum_{\tau}^{t-1} \frac{\partial \log \mathbf{f}'_\tau}{\partial \boldsymbol{\lambda}} (\boldsymbol{\xi}_{\tau|t} - \boldsymbol{\xi}_{\tau|t-1}), \tag{32}$$

where the smoothed probabilities $\boldsymbol{\xi}_{\tau|t}$ are obtained by running the Kim (1994) smoother starting from filtered probability $\boldsymbol{\xi}_{t|t}$. The partial derivatives w.r.t. to shape parameters of regime m, $\lambda_m = \{\mathbf{c}_m, \Phi_m, \Sigma_m\}$ with $m \in \{1, 2, \ldots, M\}$, derived using matrix identities (Petersen, Pedersen et al., 2008), are

$$\frac{\partial \log \mathbf{f}'_t}{\partial \mathbf{c}_m} = \Sigma^{-1}(\mathbf{y}_t - \mathbf{u}_{mt}), \tag{33}$$

$$\frac{\partial \log \mathbf{f}'_t}{\partial \boldsymbol{\Phi}_m} = \Sigma^{-1}(\mathbf{y}_t - \mathbf{u}_{mt}) \mathbf{y}'_{t-1} \tag{34}$$

and

$$
\begin{aligned}
\frac{\partial \log \mathbf{f}_t'}{\partial \boldsymbol{\Sigma}_m} = &-\frac{1}{2}\left(2(\boldsymbol{\Sigma}_m^{-1} - \boldsymbol{\Sigma}_m^{-1}\left(\boldsymbol{y}_t - \mathbf{u}_{mt}\right)\left(\boldsymbol{y}_t - \mathbf{u}_{mt}\right)'\boldsymbol{\Sigma}_m^{-1})\right) \\
&- \left(\operatorname{diag}(\boldsymbol{\Sigma}_m^{-1}) - \operatorname{diag}(\boldsymbol{\Sigma}_m^{-1}\left(\boldsymbol{y}_t - \mathbf{u}_{mt}\right)\left(\boldsymbol{y}_t - \mathbf{u}_{mt}\right)'\boldsymbol{\Sigma}_m^{-1})\right).
\end{aligned} \tag{35}
$$

For the parameters $\boldsymbol{\kappa} = \{\mathrm{P}, \boldsymbol{\rho}\}$, the conditional score defined in equation 30 becomes

$$
\begin{aligned}
&\sum_{s_t=0}^{1}\sum_{s_{t-1}=0}^{1} \frac{\partial \log P(s_t|s_{t-1}, \boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} P(s_t, s_{t-1}|\mathbf{y}_t, \boldsymbol{\theta}) \\
&+ \sum_{\tau=2}^{t-1}\sum_{s_t=0}^{1}\sum_{s_{t-1}=0}^{1} \frac{\partial \log P(s_\tau|s_{\tau-1}, \boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} P(s_\tau, s_{\tau-1}|\mathbf{y}_t, \boldsymbol{\theta}) - (P(s_1|\mathbf{y}_{t-1}, \boldsymbol{\theta})) \\
&+ \sum_{s_1=0}^{1} \frac{\partial \log P(s_1, \boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} (P(s_1|\mathbf{y}_t, \boldsymbol{\theta}) - P(s_1|\mathbf{y}_{t-1}, \boldsymbol{\theta})).
\end{aligned} \tag{36}
$$

The conditional probabilities $P(s_t, s_{t-1}|\mathbf{y}_t, \boldsymbol{\theta})$ can be computed in a similar way as the terms $p^*_{ij}(t)$, defined in equation 16. The partial derivatives w.r.t to $\boldsymbol{\rho}$ are not included as the regularity condition, that the parameter should not lie at their boundary, is not met.

4.2 The ICAPM

At the base of this derivation is the log-linear approximation of the returns based on the investor's expectation of future dividend growth and stock returns, introduced by Campbell and Shiller (1988). This approximation is given by

$$
r_{t+1} - \mathrm{E}_t[r_{t+1}] \approx (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty} \rho^j r_{t+1+j} = N_{CF,t+1} - N_{dr,t+1}, \tag{37}
$$

where r_t the log asset return at time t, d_t the log dividend paid by the firm of the corresponding asset, E_t is the rational expectation at time t and ρ represents a constant annual discount coefficient. This coefficient reflects the investor's preference for receiving cash sooner rather than later due to risk and the opportunity cost of investing. Campbell and Vuolteenaho (2004) show that the ICAPM is robust for annual discount factors ranging from 0.93 to 0.97. The annual discount rate is set to 0.95. For monthly data, this means that the disocunt rate is $0.95^{\frac{1}{12}} \approx 0.9957$. Δd is the change in log dividend at time t with respect to the period before $t-1$. This approximation decomposes the returns into two terms. The first term represents the news about future cash-flows at time $t+1$, denoted as $N_{cf,t+1}$. The second term represents the news about future discount rates and is denoted as $N_{dr,t+1}$ as it is the expected discounted returns.

In order to implement this variance decomposition we follow a VAR approach described by Campbell (1991). We define an $n \times 1$ vector $\mathbf{y}_{t+1}$ consisting of n variables. The first element of $\mathbf{y}_{t+1}$ is the stock return and the assumption is made that $\mathbf{y}_{t+1}$ can be described by a first order VAR model, i.e

$$\mathbf{y}_{t+1} = \mathrm{A}\mathbf{y}_t + \mathbf{w}_{t+1}. \tag{38}$$

We introduce an $n \times 1$ vector $\mathbf{e}_1$ which consists of zeros except for the first element, which is equal to 1. We can find an expression for expected returns assuming the VAR model trough

$$\mathrm{E}[r_{t+1+j}|\mathbf{y_t}] = \mathbf{e}_1' \mathrm{A}^{j+1}\mathbf{y}_t \tag{39}$$

Consequently, the discount rate news term $N_{dr,t+1}$ can be written as

$$N_{dr,t+1} = \mathbf{e}_1' \sum_{j=1}^{\infty} \rho^j \mathrm{A}^j \mathbf{w}_{t+1}, \tag{40}$$

which for convenience will be written as $\boldsymbol{\lambda}\mathbf{w}_{t+1}$ with $\boldsymbol{\lambda}$ equal to $\mathbf{e}_1'\rho\mathrm{A}(\mathrm{I}-\rho\mathrm{A})^{-1}$.The expression can be derived from equation 40 as it is a convergent series. This result and the variance decomposition formulated by equation 37 in turn allow us to find an expression for the cash-flow news term. $N_{cf,t+1}$ can be formulated as

$$N_{cf,t+1} = (\mathbf{e}_1' + \boldsymbol{\lambda})\mathbf{w}_{t+1}. \tag{41}$$

Campbell and Vuolteenaho (2004) builds upon these formulations of the cash-flow and discount rate news to construct the cash-flow and discount rate beta. The cash-flow beta is defined as

$$\beta_{i,cf} = \frac{\mathrm{Cov}(r_{i,t}, N_{cf,t})}{\mathrm{Var}(r^e_{M,t} - \mathrm{E}_{t-1}[r^e_{M,t}])}, \tag{42}$$

and the discount rate beta is defined as

$$\beta_{i,dr} = \frac{\mathrm{Cov}(r_{i,t}, -N_{dr,t})}{\mathrm{Var}(r^e_{M,t} - \mathrm{E}_{t-1}[r^e_{M,t}])}. \tag{43}$$

These betas are formulated such that their sum is equal to the market beta, i.e. $\beta_{i,cf} + \beta_{i,dr} = \beta_{i,M}$.

4.3 The ICAPM-MS

To generalize the model such that the agents incorporate the information of possible regimes, we follow Bianchi (2016) and Bianchi (2020).

First, given n state variables and m regimes, we define an $nm \times 1$ vector $\mathbf{q}_t = [\mathbf{q}_t^1, \ldots, \mathbf{q}_t^m]$. Where, $q_t^i = \mathrm{E}[\mathbf{y}_t \mathbb{1}_{S_t=i} | I_0]$. Here, $\mathbf{y}_t$ are the n state variables and $\mathbb{1}$ is an indicator function such that $\mathbb{1}_{S_t=i}$ is 1 if state $S_t = i$. We rewrite

$$q_t^i = \mathrm{E}[\mathbf{y}_t \mathbb{1}_{S_t=i} | I_0] = \mathrm{E}[\mathbf{y}_t | S_t = 1, I_0] P(S_t = i | I_0) \tag{44}$$

and denote $P(S_t = i|I_0)$ as π_t^i. The vector $\boldsymbol{\pi}_t$ contains all elements π_t^i, i,e $\boldsymbol{\pi}_t = [\pi_t^1, ..., \pi_t^m]$. This is similar to the definition of the state variable given in equation 3. $\boldsymbol{\pi}_t$ could therefore also be written as $\boldsymbol{\xi}_{t|0}$ assuming the number of regimes m is 2. For now we will continue using $\boldsymbol{\pi}_t$.

The variable $\mathbf{q}_t$ cannot be recovered by the previously discussed algorithms for $t \geq 2$ since the prediction method only forecasts the states one step ahead. Using $\mathbf{q}_0$ and the estimates of the parameters of the MSVAR model, a recursive forecasting method is formulated that predicts $\mathbf{q}_t$ multiple steps ahead, described as the law of motion of $\mathbf{q}_t$. The law of motion of $\mathbf{q}_t$ is given by

$$\begin{bmatrix} \mathbf{q}_t \\ \boldsymbol{\pi}_t \end{bmatrix} = \begin{bmatrix} \Omega & \mathrm{CP} \\ & \mathrm{P} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{t-1} \\ \boldsymbol{\pi}_{t-1}, \end{bmatrix} \tag{45}$$

Where P is the transition matrix, C is a block diagonal matrix bdiag$[\mathrm{C}_1, \ldots, \mathrm{C}_m]$ with matrix C_i a diagonal matrix with the elements of the intercepts given state i, $\mathbf{c}_i$, defined by equation 1, on the diagonal. In mathematical notation, this means that $\mathrm{C}_i = \mathrm{diag}(\mathbf{c}_i)$. The matrix Ω is defined as $\mathrm{bdiag}(\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_m)(\mathrm{P} \otimes \mathrm{I}_n)$, where $\boldsymbol{\Phi}_i$ is the matrix consisting of the autoregressive coefficients given state i, I_n is an $n \times n$ identity matrix, and operator $\otimes$ is the Kronecker product.

Similarly, $\mathbf{q}_{t+s|t}^i$ can be derived using the law of motion by noting that $\mathbf{q}_{t|t}^i = \mathbf{y}_t \boldsymbol{\pi}_{t|t}^i$. $\boldsymbol{\pi}_{t|t}^i$ is the smoothed estimate of the states. It is this last result, $\mathbf{q}_{t+s|t}^i$, that is important to generalize the standard ICAPM. To understand this, we revisit the formulation of the agents' expected stock returns at time $t + 1 + j$ using all available information op until time t, given by equation 39. The expected returns are derived using the VAR(1) model that does not account for multiple regimes. By understanding that $\mathbf{q}_{t+s|t}^i = \mathrm{E}[\mathbf{y}_t \mathbb{1}_{S_t=i} | I_t]$ and is a $mn \times 1$ vector, we derive

$$\mathrm{E}[\mathbf{y}_{t+s} | I_t] = \mathrm{W} \mathbf{q}_{t+s|t}^i, \tag{46}$$

with W being an $n \times mn$ matrix consisting of m identity matrices I_n placed next to each other.

The first element of $\mathbf{y}_t$ should be the excess market returns. With this formulation, the returns are forecasted while considering regime switching as well. We can now generalize equation 39, which reads

$$\mathrm{E}_t[r_{t+1+j}] = \mathbf{e_1}'\mathrm{W}\mathbf{q}^i_{t+1+j|t}. \tag{47}$$

Using this expression, the generalized definitions of the cash-flow and discount rate news can be derived in a similar way as previously discussed. The discount rate news is given by

$$N_{dr,t} = \mathbf{e_1}'\mathrm{W}[\boldsymbol{\lambda}^{\mathbf{q}}\mathrm{v}^{\mathbf{q},t} + \boldsymbol{\lambda}^{\boldsymbol{\pi}}\mathrm{v}^{\boldsymbol{\pi},t}] \tag{48}$$

and

$$N_{cf,t} = \mathbf{e_1}'\mathrm{W}[(\mathrm{I}_r + \boldsymbol{\lambda}^{\mathbf{q}})\mathrm{v}^{\mathbf{q},t} + \boldsymbol{\lambda}^{\boldsymbol{\pi}}\mathrm{v}^{\boldsymbol{\pi},t}], \tag{49}$$

with $\boldsymbol{\lambda}^{\mathbf{q}} = \rho\Omega(\mathrm{I}_r - \rho\Omega)^{-1}$ and $\boldsymbol{\lambda}^{\mathbf{q}} = \rho\mathrm{CP}(\mathrm{I}_r - \rho\Omega)^{-1}(\mathrm{I}_r - \rho\mathrm{P})^{-1}$. Furthermore, $\mathrm{v}^{\mathrm{q},t} = \mathrm{q}_{t+1|t+1} - \mathrm{q}_{t+1|t}$ and $\mathrm{v}^{\boldsymbol{\pi},t}\boldsymbol{\pi}_{t+1|t+1} - \boldsymbol{\pi}_{t+1|t}$.

4.4 Asset pricing tests

The Fama-Macbeth two-pass procedure consists of the following regressions for the ICAPM (Fama & MacBeth, 1973). First, the ICAPM betas are estimated via a timeseries regression of the excess returns of each of the portfolio returns i.

$$r^e_i = \lambda_0 + \beta_{i,dr}N_{dr} + \beta_{i,cf}N_{cf} + \boldsymbol{\epsilon}_i, \tag{50}$$

The ability of the model to explain the cross-sectional variation in average excess return of the portfolios, $\bar{r^e}$ can be evaluated via

$$\bar{\mathbf{r}}^e = \lambda_0 + \lambda_1\boldsymbol{\beta}_{dr} + \lambda_2\boldsymbol{\beta}_{cf} + \boldsymbol{\alpha}, \tag{51}$$

with excess returns $\bar{r}^e$, and estimated betas for portfolios $1, \ldots, N$. λ_0 is an intercept and λ_1 and λ_2 risk premiums for their corresponding risk factors. Given that the dependent variable is the excess market returns, it is expected that λ_0 is zero, as riskless returns should be theoretically equal to the risk-free rate.

One method to evaluate the performance is to test whether the error terms α in equation 51 are jointly zero. Heteroskedasticity in the average returns of the portfolios is not uncommon. An

OLS regression that assumes i.i.d error terms can thus produce spurious results. To account for heteroskedasticity, the cross-sectional regression is conducted through GLS. Using this method, the portfolios are essentially repackaged. The GLS test relies on the following result. In the limit, the estimate $\hat{\alpha}$ converges to a normal distribution. Consequently,

$$J = \hat{\boldsymbol{\alpha}}' \text{Var}(\hat{\boldsymbol{\alpha}})^{-1} \hat{\boldsymbol{\alpha}} \sim \chi^2_{N-k}, \tag{52}$$

with N portfolios and k factors. Since the variance of $\hat{\boldsymbol{\alpha}}$ cannot be inverted, we rewrite $\hat{\boldsymbol{\alpha}}' \text{Var}(\hat{\boldsymbol{\alpha}})^{-1} \hat{\boldsymbol{\alpha}}$ as $\hat{\boldsymbol{\alpha}}' \tilde{\Sigma}^{-1} \hat{\alpha}$. With covariance matrix $\tilde{\Sigma}$ defined as $\text{E}[\tilde{\mathbf{e}}, \tilde{\mathbf{e}}']$ and $[\tilde{\mathbf{e}}] = [\mathbf{e}_1, \ldots, \mathbf{e}_N]'$. After adding the Shanken correction term, the J statistic can be determined by

$$J = T(1 + \lambda' \tilde{\Sigma}_f^{-1} \hat{\boldsymbol{\lambda}})^{-1} \hat{\boldsymbol{\alpha}}' \tilde{\Sigma}^{-1} \hat{\boldsymbol{\alpha}}. \tag{53}$$

Here, λ denotes a vector of risk premiums and α denotes the error terms of the second pass regression described by equation 51. T is the sample size. Using the result, by evaluating this J-statistic formulated by Cochrane (2009), it can be determined whether the error terms $\boldsymbol{\alpha}$ are significantly close to zero.

The GLS R^2, related to the J-statistic, defined as

$$R^2 = 1 - \frac{\boldsymbol{\alpha}' \Sigma^{-1} \boldsymbol{\alpha}}{(\bar{r}^e - \tilde{\lambda}_0 \boldsymbol{\iota})' \Sigma^{-1} (\bar{r}^e - \tilde{\lambda}_0 \boldsymbol{\iota})} \tag{54}$$

can be interpreted as the distance of the factors' mimicking portfolio with the mean-variance boundary, as is shown by Lewellen, Nagel and Shanken (2010). A GLS R^2 of 1 would mean that the mimicking portfolio is on the mean-variance boundary. Here, $\tilde{\lambda_0}$ is the intercept of the same GLS regression on a constant.

Chapter 5

Results

5.1 Stationarity

For a standard VAR model, which is used as a benchmark model, assume stationary variables. Therefore, variables deemed nonstationary are differenced in both the VAR and MSVAR models, as the VAR model relies on the state variables being stationary. While the MSVAR model does not necessarily rely on individual variables to be stationary, this way the models consist of the same specification of the state variables. The results of the ADF test are reported in table 1 in appendix A. The results strongly suggest the presence of a unit root for the D/P- and P/E ratio. To avoid model misspecification, all observations of the D/P ratio differenced by the moving average of the 6 previous months. There is no consensus on the stationarity of the P/E ratio, particularly in light of more recent data. To stay consistent with the literature that use similar methodology to construct the ICAPM, the P/E ratio is not differenced. Using the same reasoning, the small value spread is not differenced.

5.2 The MSVAR model

The MSVAR model is estimated with the excess market returns, term yield spread, small value spread, price-to-earning ratio and the dividend-to-price ratio as state variables, with monthly data from 07-1928 until 12-2022 (1334 observations). Four regimes are assumed to exist, with two level regimes and two variance regimes with the level and variance regimes independent from each other. Figure 1 shows the estimated smoothed probabilities that level regime 1 prevails. This regime can be considered as a recession regime given the overlap with the NBER recessions, depicted as shaded areas in the figure. Evaluating regime-specific expectations reported in table 3a, the recession regime can be characterized by negative excess market returns, and an elevated term yield spread as well as increased small value spread.

Table 1 shows the estimated level parameters for the corresponding regimes. It can be observed from this table that autoregressive coefficients of the excess market returns (r_M) are substantially large for the recession regime with respect to the stable regime (level regime 2). The interdependence of the market returns and the lagged state variables thus increases during a recession regime, which has implications for the construction of the discount rate- and cash-flow news terms, given the way they are formulated in equation 40 and 41. This lagged relationship between the market returns and the other variables is not captured by the regular VAR model, which parameters are shown in table 1 in appendix B. In particular, the AR coefficient of the D/P ratio (DP) is much larger for the market returns in recession regimes and the lagged small value spread (VS) is now negatively correlated with the market returns. It should be noted that the coefficients of the D/P ratio and the P/E ratio are less interpretable in isolation, due to them being influenced by common factors by definition. This is also indicated by the large contemporaneous correlation shown in table 3 in appendix B.

It is important to consider the change in persistence of each state variable, as persistent variables are assigned larger weights in the predicting process. Although the persistence is quite stable across the regimes, TY, VS and PE are highly persistent during the stable regime but less persistent whenever the recession regime prevails.

Table 2 reports the correlation of VAR and MSVAR estimates of the news terms. There is a strong correlation between the market news terms, but the correlations among the decomposed news terms are less pronounced. This implies that a relatively similar market shock is decomposed differently by investors that consider recession and stable regimes and those who do not. Figure 2 illustrates this by showing the difference between the ICAPM factors constructed via the VAR model and the factors constructed using the MSVAR model. Here, the MSVAR factors are subtracted from the VAR factors. The negative of the discount rate news is taken such that a positive value corresponds to 'positive' discount rate news. The shaded areas represent the periods that the recession regimes prevail (during these periods the smoothed probabilities of the recession regime exceed 0.5).

It can be observed that, during the stable regime, the difference in the (negative) discount rate shocks is systematically positive whereas the difference in cash-flow shocks is negative. The opposite applies during recession regimes. This is partly caused by the frequent positive shocks of the market during stable regimes and negative shocks during recession regimes in the VAR model, which can be observed in figure 1 in in appendix C. This might be a consequence of using the information of both stable and recession regimes to forecast market returns, while not making inferences of the current and future regimes. From the standard deviations and correlations in table 2 for the VAR model, it can be observed that the discount rate news terms have a large variation with respect to the cash-flow news term and is strongly correlated to the market shock terms. A large proportion of the market shocks can therefore be attributed to discount rate news, which is consistent with Campbell (1991) and Campbell and Vuolteenaho (2004). This leads to a reformulation of the previous argument: investors that do not consider regimes discount their future earnings by too much during stable regimes.

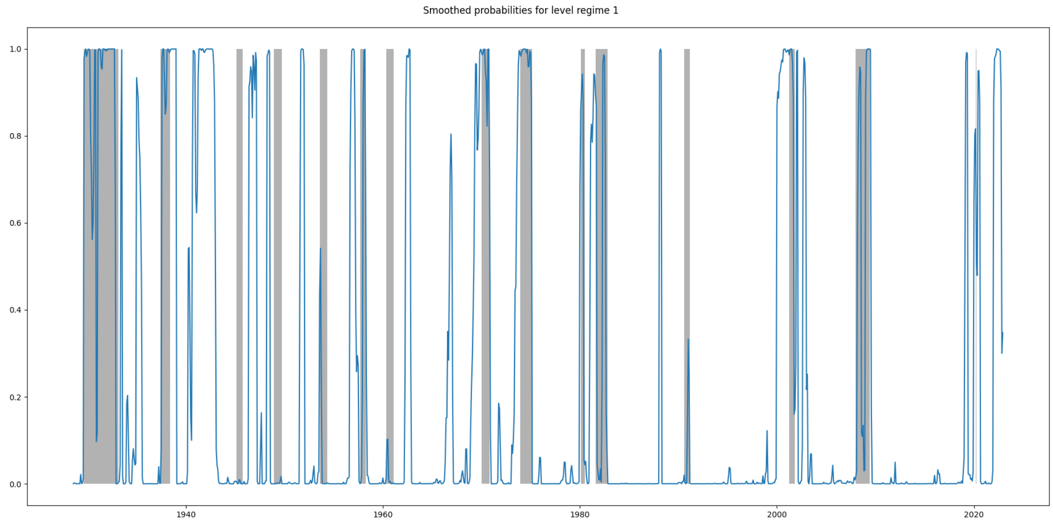

Figure 1: Smoothed probabilities of the crisis level regime estimated via an MSVAR model assuming two level regimes and two variance regimes. The state variables are the excess market returns, term yield spread, small value spread, price to earning ratio and the dividend to price ratio, with monthly data from 07-1928 until 12-2022 (1334 observations). The shaded areas correspond to NBER recessions. The excess market returns are scaled by a factor $\frac{1}{100}$ in this table

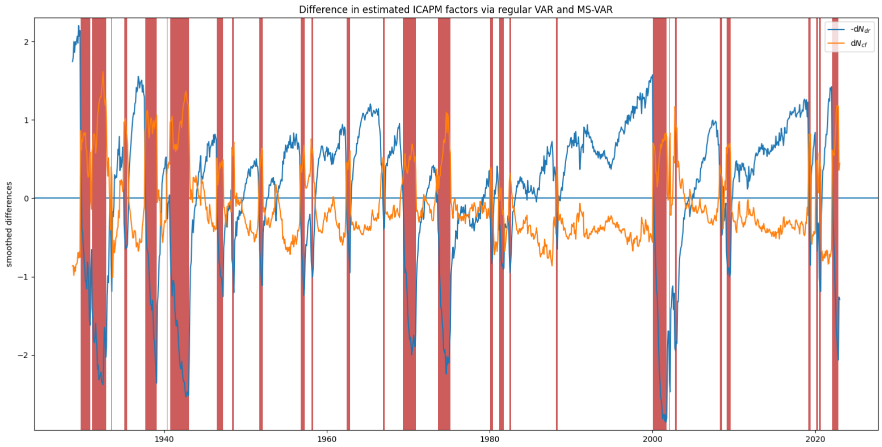

Figure 2: Differences between the decomposed market factors constructed via the VAR model and the factors constructed using the MSVAR model, in percentage points. -dN_{dr} and dN_{cf} corresponds to the difference between the estimated (negative) discount rate and cash-flow news term between the models, respectively. The MSVAR model factors have been subtracted from the VAR-model factors. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5. The estimates range from 07-1928 until 12-2022

Table 1: Estimations of the level parameters MSVAR

Regime 1	c	$r^e_{M,t}$	TY_t	VS_t	DP_t	PE_t
$r_{M,t+1}$	0.070	0.173	0.007	−0.018	0.119	−0.020
TY_{t+1}	0.676	0.020	0.940	−0.039	0.198	−0.173
VS_{t+1}	0.001	0.055	0.000	0.999	0.040	0.000
DP_{t+1}	−0.050	−0.581	−0.004	0.007	0.692	0.015
PE_{t+1}	0.029	0.659	0.002	−0.009	0.124	0.989
Regime 2						
$r^e_{M,t+1}$	0.003	−0.015	−0.001	0.011	−0.002	−0.003
TY_{t+1}	−0.103	0.202	0.969	0.065	−0.369	0.005
VS_{t+1}	0.002	0.066	0.000	0.996	−0.029	0.001
DP_{t+1}	−0.003	−0.361	0.000	0.002	0.701	0.000
PE_{t+1}	−0.004	0.385	0.000	0.006	0.044	0.998

Note: Estimated parameters of the level variables of the MSVAR model from 07-1928 until 12-2022. Level regime 1 corresponds to the recession regime and level regime 2 to the stable regime.

Table 2: correlations and standard deviations of the estimated news terms

	MSVAR			VAR			
	$N_{dr,ms}$	$N_{cf,ms}$	$N_{mkt,ms}$		N_{dr}	N_{cf}	N_{mkt}
$N_{dr,ms}$	0.0382			N_{dr}	0.0417		
$N_{cf,ms}$	0.0317	0.0367		N_{cf}	-0.2330	0.0249	
$N_{mkt,ms}$	-0.7105	0.6808	0.0522	N_{mkt}	-0.8907	0.6497	0.0533
$\text{corr}(N_{dr}, N_{dr,ms})$	0.7734						
$\text{corr}(N_{cf}, N_{cf,ms})$	0.7833						
$\text{corr}(N_{mkt}, N_{mkt,ms})$	0.9677						

Note: correlations and standard deviations of the discount rate news N_{dr}, cash-flow news N_{cf}, and market news N_{mkt}, VAR estimates on the right-hand side and the news terms MSVAR estimates on the left-hand side ($N_{dr,ms}$, $N_{cf,ms}$ and $N_{mkt,ms}$). The off-diagonal elements correspond to the correlations and on the diagonal the standard deviations of each news term is given. underneath the correlation matrices the correlations of the VAR estimates of the news terms and their counterparts estimated by the MSVAR models are given.

Table 3: Additional MSVAR estimation results

(a) Level regime-specific expectations of the state variables

	Regime 1	Regime 2
r_m^e	-0.009	0.010
TY	1.283	1.495
VS	1.881	1.679
DP	0.020	-0.008
PE	2.752	2.873

(b) Estimated transition matrix P of the MSVAR model

P	$(1,1)_t$	$(1,2)_t$	$(2,1)_t$	$(2,2)_t$
$(1,1)_{t+1}$	0.683	0.273	0.024	0.000
$(1,2)_{t+1}$	0.174	0.611	0.021	0.036
$(2,1)_{t+1}$	0.050	0.000	0.379	0.136
$(2,2)_{t+1}$	0.093	0.116	0.576	0.828

Note: Subtable (a) shows the level regime-specific expectations of the state variables. The expectations are calculated as the sum of all data points of each variable, weighted by the smoothed probabilities of the corresponding regime and subsequently divided by the total sum of the smoothed probabilities. In subtable (b), the transition matrix P of the MSVAR model is presented. The state (i, j) corresponds to the state with level regime i and variance regime j. Level regime 1 and 2 are considered the crisis and stable regime, respectively. Variance regime 1 and 2 are considered high and low volatility regimes, respectively. No standard errors for the transition matrix P are reported, as these errors were insignificant w.r.t the estimation.

The MSVAR model assumes four states, which are a combination of two level- and two variance regimes. While the the level regimes are able to be identified as recession and stable regimes, the estimated variance regimes are harder to interpret. The smoothed probabilities of variance regime 1 are shown in figure 1 in appendix B, and the estimated contemporaneous correlations and standard deviations are given in table 3 in appendix B. Variance regime 1 does seem to prevail more often during recessions, and can further be regarded as the high volatility state. In the aforementioned table, it can be observed that the volatility of each state variable is substantially higher for variance regime 1. However, a remarkable pattern can be observed in the smoothed probabilities of variance regime 1. During stable periods, variance regime 1 also prevails specifically during June every year. The cause of this regime to reoccur in June every year may simply be due to the reconstruction of the portfolios that occurs at June each year which is used to define the small value spread. This reconstruction often results in a relatively big correction in the small value spread, which is interpreted as a large shock. The relative difference in standard deviation of the small value spread is significantly larger across the regimes with respect to the volatility of the other variables. In fact, this standard deviation is more than ten times as large for the high volatility regime, while the stand deviations of the other variables are approximately twice as large.

Table 3b shows transition probabilities of the MSVAR model. As expected, the combination of the stable level regime and the low volatility regime is the most persistent regime. Notable are the large transition probabilities from a low to high volatility regime and vice versa. While this effect may be caused due to the seemingly annual transition of volatility regimes in June during the stable level regimes, the effect is also present during recession level regimes. This suggests that volatility is dynamic and fluctuates during economic downturns. Figure 2 in appendix B shows the smoothed probabilities of the recession regime in combination with the volatility regimes, highlighting the dynamics of the volatility regimes. The recession-high volatility regime is more persistent and prevails more often during earlier periods, especially during the Great Depression. While the recession-low volatility regime is inpersistent, it tends to persist for extended periods during certain phases. Again, this can be partly attributed to the identification of high volatility regimes during stable periods, which occurs due to the small value spread. If these specific periods coincide with low volatility in other state variables, the identification of the 'high volatility' regime may be perturbed.

5.3 The ICAPM-MS and asset pricing test results

Figures 3 and 4 show the performance measures of the second pass cross-sectional regression on 24 size-value sorted portfolios. The extreme small-growth portfolio is not included in these tests as it is considered an outlier in asset pricing models. For each model, there is no intercept included in the second pass. Betas are estimated from 01-1967 until 12-2022 using a rolling window of 35 years as sample. The factors of the ICAPM-MS consist of the market news components estimated by the MSVAR model whereas the standard ICAPM has the two news components estimated by the VAR model as factors. Figure 3 reports the cross-sectional GLS R^2 and figure 4 the p-values of the cross-sectional asset pricing tests. Based on the GLS R^2, an improvement in performance of the ICAPM-MS can be observed with respect to the ICAPM whenever during transitions of regimes. An improved GLS R^2 implies that a larger proportion of the cross-sectional variation in the portfolio returns is explained by the ICAPM-MS. While the overall GLS R^2 is fairly low, and the observed improvement may not be substantial, the p-values of the cross-sectional asset pricing tests show the same improvements during similar periods.

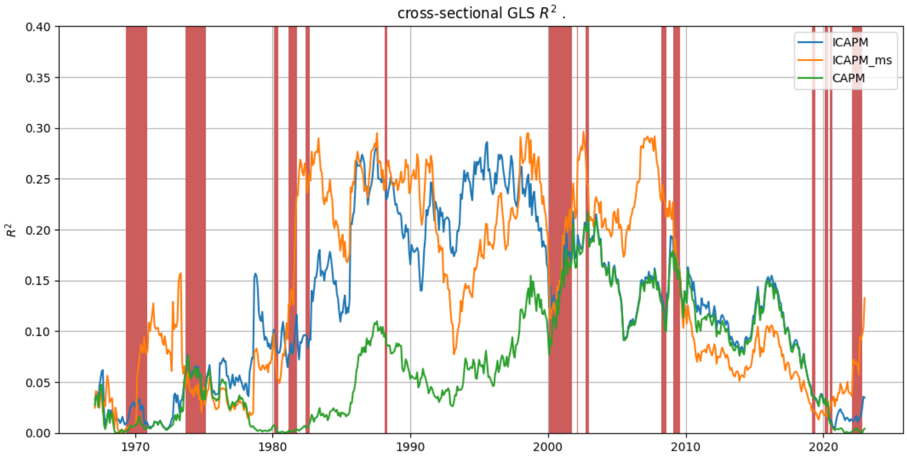

Figure 3: R^2 of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 35-year moving window from 01-1967 until 12-2022. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

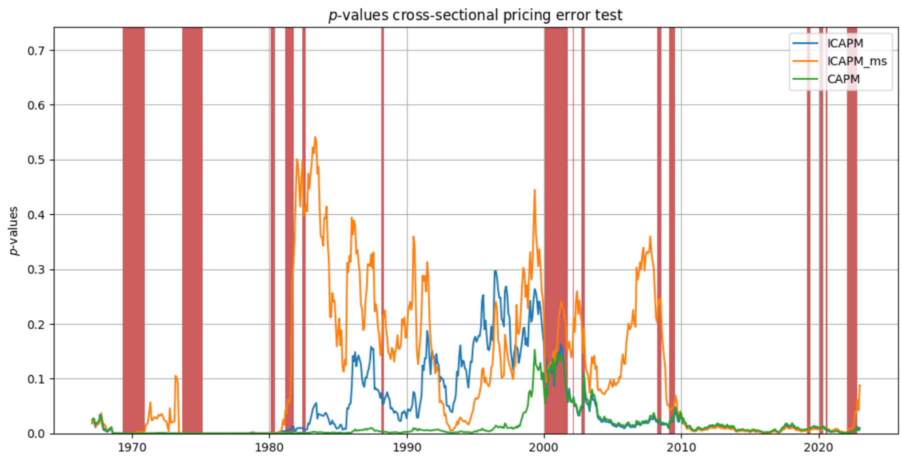

Figure 4: p-values of the asset pricing test of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 35-year moving window from 1967-01 until 2022-12. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

It is worth investigating whether there is a clear cause for the improved performance of the ICAPM-MS around the years 1975, 1983 and 2005. Figure 5 shows the progression of the difference between the ICAPM-MS betas and the ICAPM betas for each portfolio. The betas based on the factors estimated by the MSVAR model are subtracted from the betas based on the estimated factors by the VAR model. If the improved performance was due to the ICAPM-MS being able to better explain the size or value anomaly, one would expect a persistent difference between the betas, e.g. the difference in the progression of the betas of the small portfolios would be systematically positive, which is not the case. During the periods from 1980 until 1990 and from 2005 until 2010, the p-values of the asset pricing test do not reject the null that the error terms are zero for the ICAPM-MS, while rejecting the ICAPM at a 90% confidence level for most of the time. During these periods both models overestimate the aggregate returns mainly for small and value portfolios. The difference is that the positive error terms of the ICAPM-MS are not as large as those of the ICAPM. The periods can be characterized by a dip in aggregate returns, mainly for small and value portfolios, and quickly decreasing discount rate betas as well as increasing cash-flow betas for the ICAPM-MS. This results suggests that during both periods, the sensitivities to discount rate news of high earning stocks is underestimated by investors not recognizing the effect of a recession regime on returns. The sensitivity to market shocks is acknowledged but misinterpreted as sensitivity to cash-flow news. These effects are however very small, as the variation in betas is not significant

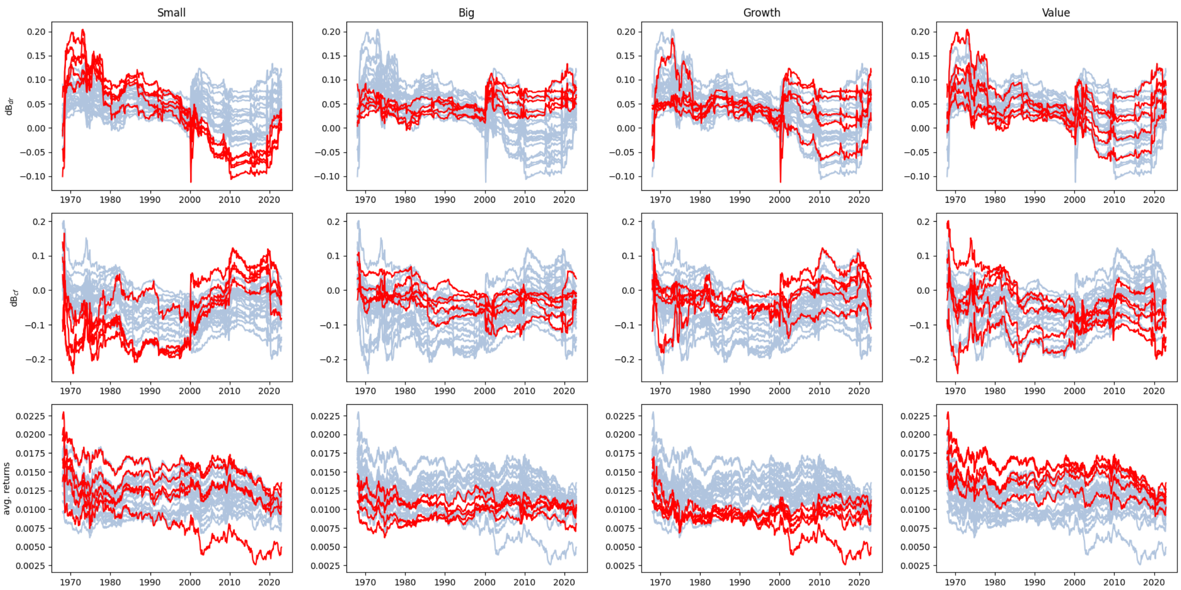

Figure 5: The first two rows show the progression of the difference between the ICAPM-MS betas and the ICAPM betas for each of the portfolios. The betas are estimated using a 35-year rolling window. The betas based on the factors estimated by the MSVAR model are subtracted from the betas based on the estimated factors by the VAR model. The third row shows the average return of each of the portfolios. In the first column, the five small portfolios are highlighted in red. In the second column the five big portfolios, in the third column the five growth portfolios and in the fourth column the five value portfolios are highlighted.

Chapter 6

Robustness tests

6.1 Asset pricing test on the full modern sample

Betas are estimated for each model for the 24 size-value sorted portfolios, based on the 'modern' sample ranging from 03-1963 until 12-2022. Each model does contain an intercept along with the betas. The estimated betas of the ICAPM-MS per portfolio are shown in table 1. In parentheses, the difference of the estimated beta with respect to the estimated betas using the ICAPM is given. For the ICAPM, the estimated betas are similar to the estimated betas by Campbell and Vuolteenaho (2004) on a modern sample to some degree. Their sample starts at 03-1963 and ends at 12-2001. The main difference in estimated betas is that in general, the discount rate betas are larger while the cash-flow betas are much smaller (in the range of 0.10). Another difference is that the discount rate betas do not show much variation across the value spectrum of the portfolios in the more updated sample.

The discount rate betas are still relatively capable in explaining the spread in returns in Campbell, Giglio, Polk and Turley (2018), where the sample ranges from 03-1963 until 4-2011. This effect can also be observed in figure 5, where growth and value portfolios do not seem to have common sensitivities to discount rate news for the most recent samples. Further notable from table 1 are the structurally smaller discount rate betas and larger cash-flow betas. As with the rolling window samples, there is no obvious difference in variation across the portfolio sorts between the estimated betas.

Table 2 reports the performance measures and estimated risk premiums of the ICAPM, ICAPM-MS and CAPM. The OLS R^2 is calculated as 1 - SS/SS_b, where SS represents the sum of squared errors, and SS_b denotes the sum of squared errors from a regression with only an intercept term. The models have an unrestricted zero-beta rate. Although the ICAPM and ICAPM-MS have high OLS R^2, the GLS R^2 does not imply that the models successfully explain the cross-sectional variation in returns. Both discount rate risk premiums are negative, which is unrealistic. Large cash-flow and small discount rate risk premiums are however expected as cash-flow shocks are deemed to be permanent and discount rate shocks transient. The estimated risk premium for cash-flow news is to some degree similar to the cash-flow news risk premium estimated via a variant of the ICAPM by Campbell et al. (2018). Large zero-beta risk premiums can be observed. Although the zero-beta estimated by the ICAPM-MS is the lowest, it still deviates significantly from the expected zero percent.(This is expected as the the excess returns of portfolios are evaluated.) Notable are the extreme and unrealistic risk premiums for the CAPM, with a large zero-beta and negative market risk premium, further illustrating the inability of the CAPM to explain asset returns in modern samples.

Table 1: Estimated betas ICAPM-MS for the modern sample

B_{dr}	Growth	2	3	4	Value	Diff.
Small	0.92(−0.13)	0.82(−0.10)	0.77(−0.08)	0.73(−0.07)	0.75(−0.08)	0.17 (−0.05)
2	0.87(−0.15)	0.79(−0.10)	0.73(−0.09)	0.70(−0.08)	0.79(−0.09)	0.08 (−0.06)
3	0.83(−0.16)	0.75(−0.10)	0.70(−0.07)	0.71(−0.08)	0.75(−0.09)	0.08 (−0.07)
4	0.78(−0.12)	0.73(−0.09)	0.71(−0.07)	0.69(−0.09)	0.74(−0.10)	0.04 (−0.01)
Big	0.67(−0.08)	0.64(−0.08)	0.62(−0.05)	0.63(−0.07)	0.68(−0.07)	+0.01(−0.01)
Diff.	0.24(−0.05)	0.18(−0.02)	0.15(−0.04)	0.10(−0.00)	0.07(−0.01)	
B_{cf}	Growth	2	3	4	Value	Diff.
Small	0.63(+0.12)	0.57(+0.10)	0.49(+0.07)	0.46(+0.06)	0.46(+0.07)	0.17 (+0.05)
2	0.66(+0.14)	0.54(+0.10)	0.50(+0.08)	0.47(+0.07)	0.52(+0.09)	0.14 (+0.04)
3	0.63(+0.14)	0.54(+0.10)	0.48(+0.07)	0.45(+0.07)	0.50(+0.09)	0.14 (+0.04)
4	0.59(+0.11)	0.52(+0.09)	0.47(+0.07)	0.46(+0.08)	0.51(+0.09)	0.08 (+0.02)
Value	0.48(+0.06)	0.46(+0.06)	0.42(+0.04)	0.43(+0.06)	0.47(+0.07)	0.06 (−0.01)
Diff.	0.16(+0.05)	0.10(+0.04)	0.09(+0.03)	0.03(−0.00)	0.23(+0.00)	

Note: Estimated discount rate betas $B_d r$ and cash-flow betas $B_c f$ via the ICAPM-MS are estimated for value-size sorted portfolios for the modern sample, ranging from 03-1963 until 12-2022. The value in brackets is the deviation of the ICAPM-MS betas with respect to the ICAPM betas. The sixth column reports the difference in estimated betas between the growth and value portfolios for each corresponding row. The values in brackets give the deviation of this difference with respect to the difference in ICAPM beta estimations. Similarly, the sixth row of the two subtables report the difference in estimated betas between the small and big portfolios.

Table 2: Performance measures and estimated risk premiums of the ICAPM-MS and benchmark models with an unrestricted zero-beta rate.

Model	ICAPM-MS	ICAPM	CAPM
GLS R^2	0.129	0.042	0.006
	(0.070,0.142)	(0.029, 0.068)	(0.003,0.034)
OLS R^2	0.507	0.406	0.011
	(0.314,0.612)	(0.290, 0.458)	(0.005, 0.213)
λ_0	0.0072 (8.8%)	0.0114 (13.8%)	0.0473 (56.8%)
	(3.4%, 12.4%)	(6.7%, 19,1%)	(33.2%, 62.5%)
λ_{cf}	0.0230 (27.8%)	0.0221 (26.5%)	-
	(23.5%, 38.0%)	(18.4%, 33.7%)	-
λ_{dr}	-0.0010 (-1.23%)	-0.0021 (-1.57%)	-
	(-4.83%, 2.01%)	(-3.98%, 0.23%)	-
λ_m	-	-	-0.0330 (-39.6%)
	-	-	(-48.0%, -19,7%)

Note: The GLS R^2, OLS R^2 and estimated risk premiums of the ICAPM, ICAPM-MS and CAPM. λ_0 corresponds to the zero-beta rate. λ_{cf}, λ_{dr} and λ_m correspond to the risk premiums of the cash-flow, discount rate and market news, respectively. The values in parentheses next to the values of the risk premiums denote the annualized risk premiums in percentage points. For each model, the zero-beta rate is unrestricted. Underneath each estimation the standard deviations are reported. The confidence intervals are constructed by by simulating 10.000 draws of the MSVAR parameters assuming a multivariate normal distribution for the estimates. Each iteration, the factor models and performance measures are re-estimated.

6.2 MSVAR specification

By adding the regime switching component to the ICAPM, the estimations of the discount rate and cash-flow news become heavily reliant on the choice of the state variables. Campbell and Vuolteenaho (2004) found that the ICAPM was robust when different state variables are added or omitted, except for the small value spread, which was essential for reproducing similar results. As shown in Bianchi (2020), if the MSVAR model assumes separate level and variance Markov chains, the small value spread has a significant influence on the optimization process. Figure 2 in appendix C, shows the smoothed probabilities that the Great Depression regime prevails. The specifications of this model are the same as the previously discussed MSVAR model, but without the P/E ratio as state variable. Analogous to the results of Bianchi (2020), this regime prevails during the Great Depression and seems to shortly prevail during the Great Recession.

Notable is the slight increase in the probability that this regime prevails in 2020. It is found that removing any of the macroeconomic variables other than the small value spread, makes the model more inclined to find the Great Depression regime.

By imposing that the level regimes and variance regimes are independent, similar probabilities as well as parameters are found. In this case the level regimes are more distinct but the variance regimes are again driven by the variance of the small value spread. While the level regimes are more distinct, the recession regime is also inpersistent and has consequently less overlap with the NBER recessions. The smoothed probabilities are shown in figure 3 in appendix C.

The computed standard deviations of the estimates can be observed in table 1 and 3 in appendix B. It can be observed that the standard deviations can be quite significant with respect to the actual values, with occasional standard deviations larger than 50% of the absolute value of the estimates. The volatility estimates do however come with relatively small standard deviations.

6.3 Window size

The sample of the excess market returns and other state variables to estimate the MSVAR model ranges from 07-1928 until 12-2022. The betas of the ICAPM are re-estimated with a 35 year rolling window starting from 01-1967 until 12-2022. From the variation in GLS R^2 over time in figure 3, it becomes clear that the performance of the ICAPM-MS is quite sensitive to the sample. The discussed improvements do persist for smaller window sizes. To illustrate this, figures 4 and 5 in appendix C show the test statistics for a 25 year rolling window. The observed improvements of the ICAPM-MS diminish for window size larger than 45 years, which is shown in figure 6 and 7. This is attributed to the convergence of estimated betas by the two models. It is likely that the betas converge due to the inclusion of many regime switches in the sample. The estimated betas by the ICAPM-MS based on a sample that includes many regime may be similar to the betas estimated by the regime-independent ICAPM. This sample size can also be deemed as the tipping point of the overall performance, as the GLS R^2 starts to fall below 0.25 during most periods. The cross-sectional test is rejected at a 90% confidence level for most periods as well.

6.4 Additional tests

The R^2 of the cross-sectional regression via OLS using the 35 year rolling window is shown in figure 8 in appendix C. As mentioned before, the results from the OLS regression should evaluated carefully given its i.i.d assumption. What can be observed is the larger R^2 around the periods 1975, 1983 and in 2005, although being marginally different. The OLS R^2 does however come with the intuitive interpretation: it measures the accuracy of the model to predict the average returns. The overall R^2 is high from 1980 until 2008, but this does not serve as proof to a high performance of the model, as Lewellen et al. (2010) points about that is it relatively easy for any factor model to obtain such a high R^2. S

Chapter 7

Conclusion

The results of asset pricing tests on the models with a rolling window sample suggest that the performance of the ICAPM-MS during recession regimes increases, with respect to the ICAPM, due to better estimating the risk factors during recession regimes. The overall GLS R^2 of the model is fairly low and the based on the GLS R^2, it is disputable whether the improvements of the MS-CAPM can be regarded as significant. The p-values of the asset pricing test do mitigate this concern, showing significant improvements during the discussed periods. A drawback of the ICAPM-MS is the dependence on the sample. While the estimated betas are relatively consistent, the performance of the model relies heavily on the window. The models' performance for large window sizes is quite low, leaving it incapable to find a persistent relation between portfolio returns and the risk factors, or explain the abnormalities in portfolio returns for longer periods.

Based on comparisons of the factors of the two models, it could be argued that projections by investors about future cash-flows were too optimistic during stable regimes and too pessimistic during recession regimes, if regimes were not considered. On the other hand, said investors would discount future cash-flows by too much during stable regimes and not enough during recessions. While this effect was persistent, no significant improvements were made in estimating the betas and finding more accurate risk premiums. In general, it can be concluded that including the possibilities of regime switching did not lead to a better representation of investor behavior.

Nevertheless, there was indication that ICAPM-MS outperformed the ICAPM during specific periods. The improved performances occurred after the transition from a recession regime to a stable regime. This may be due to more positive expectations of the discount rate of investors that consider the regimes during these periods. This would imply that investors do in fact consider the transition to a stable or expansion regime at the end of a recession.

The improved performance of the ICAPM-MS was not due to its ability to better explain the size or value anomaly. Rather, the improvement were largely attributed to its enhanced capability to predict average portfolio returns in general, during the specified periods. These periods are relatively short the improvements are not observed for the majority of time. This concept of short-lived improvements is also discussed in (L. E. Farmer, Schmidt & Timmermann, 2023), who found evidence for local stock predictability. Similarly, these periods of predictability of stock returns accur briefly.

An important reason for the low general performance of the ICAPM-MS can be attributed to the fact that the model is quite a modest extension of the CAPM. The discount rate and cash-flow news terms are defined by Campbell (1991) via a log-linear approximation based on future returns and dividends. This method is parsimonious as well as interpretable and the results show a significant improvement of the model with respect to the CAPM. This shows that distinguishing between the two sources of risk via this method is of importance. However, the ICAPM with the inclusion decomposition based on an MSVAR model specified as in this paper does not sufficiently explain the abnormalities in portfolio returns. One of the issues of the ICAPM-MS was the unusual estimated variance regimes.

These regimes were primarily identified through the small value spread, overshadowing identification based on volatility clustering of returns. A variant that does a better job in capturing stochastic volatility, is the extended three factor ICAPM by Campbell et al. (2018). Similar to the two factor ICAPM, the factors are derived based on the SDF of a long-term investor, but in a setting with conditional heteroskedasticity. An additional risk factor is included in the model, which represents news about volatility. They find that this additional factor helps explaining the spread excess returns in value sorted portfolios, which did not seem to be explained by the regime dependent cash-flow and discount rate news terms. This method allows for a more flexible representation of the stochastic volatilty. The three factor ICAPM might be more adept at capturing complex patterns of volatility clustering, without the assumption of recession and stable regimes.

Another approach that can improve the overall performance of the model, is to include other factors that may contribute in explaining abnormalities in returns, which the decomposed market factors did not account for. For example, Khan (2008) constructs a four factor model consisting of the decomposed market news terms as well as the High-Minus-Low and Small-Minus-Big factors, and shows improved performance of the model in explaining abnormal returns in accrual-size sorted portfolios compared to other benchmark models. Since the main objective of this research was to compare the ICAPM-MS to the ICAPM, adding additional factors is not considered.

References

Ang, A. & Timmermann, A. (2012). Regime changes and financial markets. *Annu. Rev. Financ. Econ.*, *4*(1), 313–337.

Bianchi, F. (2016). Methods for measuring expectations and uncertainty in markov-switching models. *Journal of Econometrics*, *190*(1), 79–99.

Bianchi, F. (2020). The great depression and the great recession: A view from financial markets. *Journal of Monetary Economics*, *114*(C), 240–261.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal*, *101*(405), 157–179.

Campbell, J. Y., Giglio, S., Polk, C. & Turley, R. (2018). An intertemporal capm with stochastic volatility. *Journal of Financial Economics*, *128*(2), 207–233.

Campbell, J. Y. & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*(3), 195–228.

Campbell, J. Y. & Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, *94*(5), 1249–1275.

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

Costa, O. L. d. V. & Dufour, F. (2005). On the ergodic decomposition for a class of markov chains. *Stochastic processes and their applications*, *115*(3), 401–415.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, *39*(1), 1–22.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Fama, E. F. & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, *22*(1), 3–25.

Fama, E. F. & French, K. R. (1992). The cross-section of expected stock returns. *The Journal* *of Finance*, *47*(2), 427–465.

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

Farmer, L. E., Schmidt, L. & Timmermann, A. (2023). Pockets of predictability. *The Journal of Finance*, *78*(3), 1279–1341.

Farmer, R. E., Waggoner, D. F. & Zha, T. (2009). Understanding markov-switching rational expectations models. *Journal of Economic theory*, *144*(5), 1849–1867.

French, K. R. (2012). Kenneth r. french-data library. *Tuck-MBA program web server. http://mba. tuck. dartmouth. edu/pages/faculty/ken. french/data_library. html (accessed October 20, 2010).*

Gordon, S. & St-Amour, P. (2000). A preference regime model of bull and bear markets. *the American Economic Review*, *90*(4), 1019–1033.

Guidolin, M. & Ono, S. (2006). Are the dynamic linkages between the macroeconomy and asset prices time-varying? *Journal of Economics and Business*, *58*(5-6), 480–518.

Guidolin, M. & Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, *31*(11), 3503–3544.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business. *Econometrica*, *57*(2), 357–384.

Hamilton, J. D. (1993). Estimation, inference and forecasting of time series subject to changes in regime. *Handbook of statistics*, *11*, 231–260.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis* (pp. 202–209). Springer.

Khan, M. (2008). Are accruals mispriced? evidence from tests of an intertemporal capital asset pricing model. *Journal of Accounting and Economics*, *45*(1), 55–77.

Kim, C.-J. (1994). Dynamic linear models with markov-switching. *Journal of econometrics*, *60*(1-2), 1–22.

Kole, E. (2019). Markov switching models: An example for a stock market index. *Available at SSRN 3398954*.

Kole, E. & Van Dijk, D. (2017). How to identify and forecast bull and bear markets? *Journal of Applied Econometrics*, *32*(1), 120–139.

Kole, E. & van Dijk, D. (2023). Moments, shocks and spillovers in markov-switching var models. *Journal of Econometrics*, *236*(2), 105474.

Krolzig, H.-M. (2013). *Markov-switching vector autoregressions: Modelling, statistical inference, and application to business cycle analysis* (Vol. 454). Springer Science & Business Media.

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

Maio, P. (2013). Intertemporal capm with conditioning variables. *Management Science*, *59*(1), 122–141.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867–887.

Nagel, S. (2013). Empirical cross-sectional asset pricing. *Annu. Rev. Financ. Econ.*, *5*(1), 167–199.

Park, C. (2010). When does the dividend–price ratio predict stock returns? *Journal of Empirical Finance*, *17*(1), 81–101.

Petersen, K. B., Pedersen, M. S. et al. (2008). The matrix cookbook. *Technical University of Denmark*, *7*(15), 510.

Shiller, R. J. (1996). Price-earnings ratios as forecasters of returns: The stock market outlook in 1996. *Unpublished manuscript, Yale University*.

Tong, H. (1978). On a threshold model in pattern recognition and signal processing, ed. *CH Chen, Amsterdam: Sijhoff & Noordhoff*.

Appendix A

Stationarity state variables

Table 1: ADF test results variables

	ADF (p-value)
r_M^e	-8.52 ($<$ 0.001)
TY	-5.02 ($<$ 0.001)
VS	-2.42 (0.14)
DP	-1.22 (0.67)
PE	-1.84 (0.36)

Note: Results of the Augmented Dicky-Fuller test along with the p-value of the null hypothesis that a unit root is present. Tests are conducted on the excess market returns (r_M^e), term yield spread (TY), small values spread (VS), the D/P ratio (DP), and the P/E ratio (PE). The null is rejected for the excess market returns and term yield spread at a 1% confidence level and for the small value spread at a 15% confidence level. The null is not rejected for the D/P ratio and the P/E ratio at any reasonable confidence level.

Appendix B

Estimates VAR models

Table 1: Estimated parameters of the VAR model

(a) Estimated level parameters

	c	$r^e_{M,t}$	TY_t	VS_t	DP_t	PE_t
$r^e_{M,t+1}$	0.034	0.113	0.002	−0.004	0.036	−0.009
TY_{t+1}	0.079	0.000	0.939	0.070	−0.020	−0.039
VS_{t+1}	0.015	0.103	0.000	0.989	−0.029	0.001
DP_{t+1}	−0.019	−0.478	−0.001	0.003	0.689	0.006
PE_{t+1}	0.013	0.534	0.002	−0.006	0.065	0.997

(b) Estimated contemporaneous correlations and variances

	$r^e_{M,t+1}$	TY_{t+1}	VS_{t+1}	DP_{t+1}	PE_{t+1}
$r^e_{M,t+1}$	0.053				
TY_{t+1}	−0.026	0.343			
VS_{t+1}	0.076	−0.003	0.050		
DP_{t+1}	−0.722	0.032	−0.057	0.030	
PE_{t+1}	0.769	−0.049	0.022	−0.914	0.035

Subtable (a) shows the level parameter estimates of the VAR model. In subtable (b), the off-diagonal elements are the estimated contemporaneous correlations of the variables. The diagonal elements are the standard deviation estimates.

Table 2: Estimated parameters of the level variables of the MSVAR model.

Regime 1	c	$r^e_{M,t}$	TY_t	VS_t	DP_t	PE_t
$r_{M,t+1}$	0.070	0.173	0.007	−0.018	0.119	−0.020
	(0.057), (0.036)	(0.011), (0.001)	(0.014), (0.009)	(0.011), (0.014)	(0.047), (0.061)	(0.012), (0.007)
TY_{t+1}	0.676	0.020	0.940	−0.039	0.198	−0.173
	(0.598), (0.348)	(0.016), (0.003)	(0.016), (0.009)	(0.021), (0.002)	(0.092), (0.006)	(0.073), (0.086)
VS_{t+1}	0.001	0.055	0.000	0.999	0.040	0.000
	(0.028), (0.006)	(0.021), (0.019)	(0.004), (0.001)	(0.004), (0.006)	(0.003), (0.002)	(0.002), (0.001)
DP_{t+1}	−0.050	−0.581	−0.004	0.007	0.692	0.015
	(0.026), (0.019)	(0.141), (0.096)	(0.023), (0.008)	(0.002), (0.005)	(0.014), (0.006)	(0.005), (0.007)
PE_{t+1}	0.029	0.659	0.002	−0.009	0.124	0.989
	(0.042), (0.023)	(0.198), (0.016)	(0.003), (0.006)	(0.047), (0.043)	(0.010), (0.204)	(0.001), (0.000)
Regime 2						
$r^e_{M,t+1}$	0.003	−0.015	−0.001	0.011	−0.002	−0.003
	(0.025), (0.001)	(0.017), (0.040)	(0.000), (0.001)	(0.015), (0.006)	(0.003), (0.005)	(0.021), (0.013)
TY_{t+1}	−0.103	(0.202)	0.969	0.065	−0.369	0.005
	(0.037), (0.012)	(0.017), (0.070)	(0.038), (0.046)	(0.014), (0.022)	(0.301), (0.187)	(0.005), (0.008)
VS_{t+1}	0.002	0.066	0.000	0.996	−0.029	0.001
	(0.070), (0.003)	(0.035), (0.042)	(0.000), (0.000)	(0.005), (0.003)	(0.011), (0.021)	(0.000), (0.000)
DP_{t+1}	−0.003	−0.361	0.000	0.002	0.701	0.000
	(0.011), (0.010)	(0.142), (0.221)	(0.003), (0.000)	(0.001), (0.002)	(0.251), (0.395)	(0.000), (0.000)
PE_{t+1}	−0.004	0.385	0.000	0.006	0.044	0.998
	(0.252), (0.012)	(0.009), (0.167)	(0.001), (0.002)	(0.006), (0.011)	(0.016), (0.028)	(0.010), (0.149)

[1]

[1] *Note:* Estimated parameters of the level variables of the MSVAR model. Level regime 1 corresponds to the recession regime and level regime 2 to the stable regime. The standard deviations are constructed by assuming a multivariate normal distribution for the estimates. Due to the specification of a combination of two level- and two variance regimes, two standard deviations are reported. The left standard deviation represents the variation when a particular level regime coincides with the variance regime 1, while the right standard deviation corresponds to the level regime coinciding with the variance regime 2.

Table 3: Estimated contemporaneous correlations and standard deviations MSVAR model

Regime 1	$r^e_{M,t+1}$	TY_{t+1}	VS_{t+1}	DP_{t+1}	PE_{t+1}
$r^e_{M,t+1}$	0.081				
	(0.024), (0.106)				
TY_{t+1}	−0.057	0.502			
	(0.031), (0.045)	(0.114), (0.083)			
VS_{t+1}	0.108	−0.014	0.098		
	(0.044), (0.035)	(0.028), (0.012)	(0.016), (0.008)		
DP_{t+1}	−0.752	0.111	−0.055	0.045	
	(0.110), (0.079)	(0.880), (0.152)	(0.211), (0.301)	(0.009), (0.017)	
PE_{t+1}	0.769	−0.104	0.050	−0.945	0.053
	(0.641), (0.046)	(0.020), (0.010)	(0.315), (0.205)	(0.004), (0.006)	(0.012), (0.009)
Regime 2	$r^e_{M,t+1}$	TY_{t+1}	VS_{t+1}	DP_{t+1}	PE_{t+1}
$r^e_{M,t+1}$	0.037				
	(0.011), (0.008)				
TY_{t+1}	0.037	0.258			
	(0.019), (0.022)	(0.005), (0.006)			
VS_{t+1}	0.045	−0.024	0.008		
	(0.020), (0.033)	(−0.022), (0.017)	(0.002), (0.002)		
DP_{t+1}	−0.741	−0.071	−0.005	0.022	
	(0.041), (0.064)	(0.022), (0.016)	(0.011), (0.016)	(0.003), (0.020)	
PE_{t+1}	0.750	0.054	−0.007	−0.954	0.025
	(0.058), (0.419)	(0.022), (0.029)	(0.013), (0.009)	(0.616), (0.050)	(0.005), (0.007)

Note: Estimated variance parameters MSVAR model for two variance regimes. The off-diagonal elements are the estimated contemporaneous correlations of the variables. The diagonal elements are the standard deviation estimates. Underneath each estimation, the standard deviations are reported in parentheses. The standard deviations are constructed by assuming a multivariate normal distribution for the estimates. Due to the specification of a combination of two level- and two variance regimes, two standard deviations are reported. The left standard deviation represents the variation when a particular variance regime coincides with the recession level regime, while the right standard deviation corresponds to the variance regime coinciding with the stable level regime.

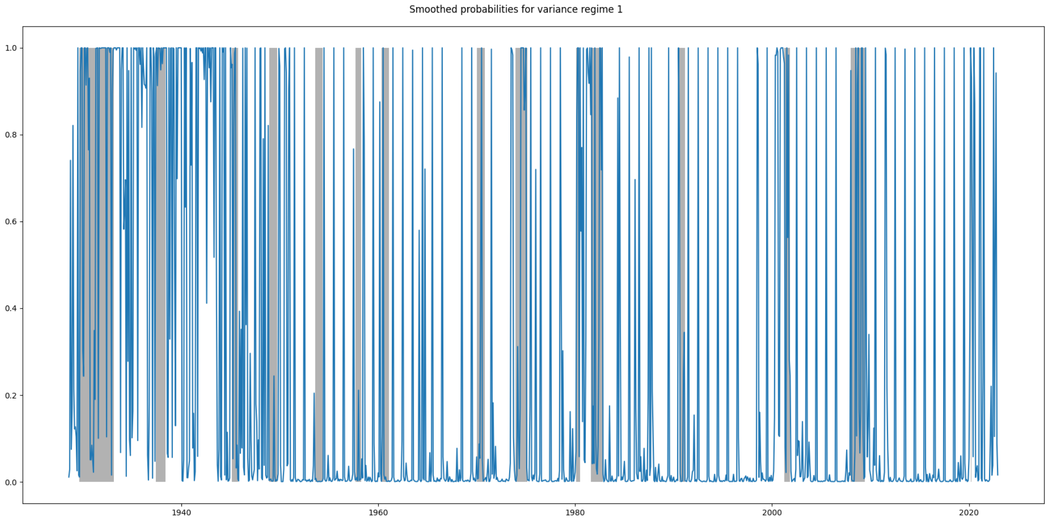

Figure 1: Smoothed probabilities of variance regime 1 estimated via an MSVAR model assuming two level regimes and two variance regimes. The state variables are the excess market returns, small value spread, price to earning ratio and the dividend to price ratio, with monthly data from 07-1928 until 12-2022 (1334 observations). The shaded areas correspond to NBER recessions.

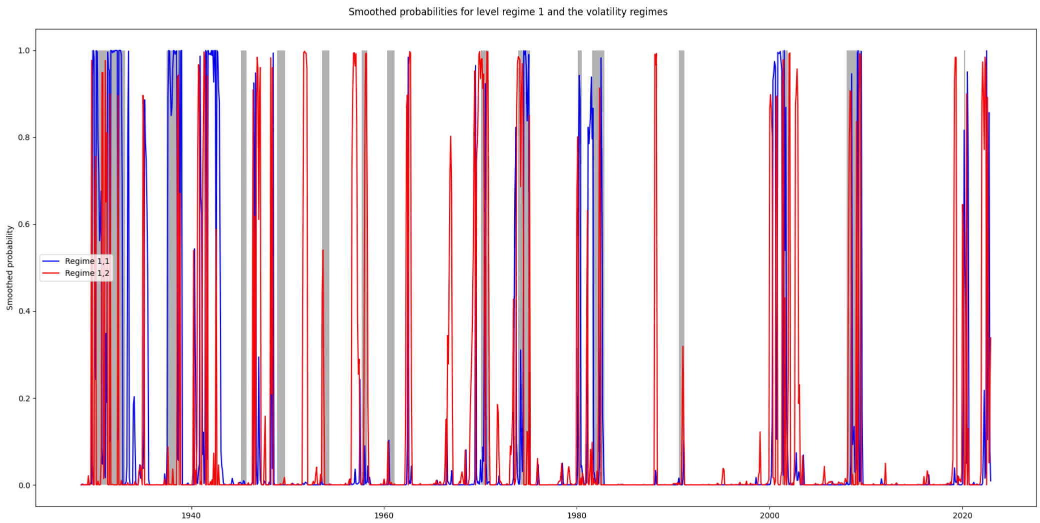

Figure 2: Smoothed probabilities of the combined recession regime (level regime 1) and the volatility regimes estimated via an MSVAR model assuming two level regimes and two variance regimes. regime 1,1 corresponds to the recession-high volatility regime while regime 1,2 corresponds to the recession-low volatility regime. The state variables are the excess market returns, small value spread, price to earning ratio and the dividend to price ratio, with monthly data from 07-1928 until 12-2022 (1334 observations). The shaded areas correspond to NBER recessions.

Appendix C

Additional tables and figures

Table 1: Estimated betas ICAPM-MS for the modern sample

B_{dr}	Growth	2	3	4	Value
Small	0.92	0.82	0.77	0.73	0.75
	(0.70, 1.05)	(0.60, 0.96)	(0.59, 0.90)	(0.54, 0.83)	(0.57, 0.83)
2	0.87	0.79	0.73	0.70	0.79
	(0.65, 0.91)	(0.58, 0.95)	(0.51, 0.81)	(0.52, 0.81)	(0.61, 0.89)
3	0.83	0.75	0.70	0.71	0.75
	(0.61, 0.92)	(0.63, 0.87)	(0.56, 0.80)	(0.64, 0.79)	(0.55, 0.85)
4	0.78	0.73	0.71	0.69	0.74
	(0.56, 0.86)	(0.50, 0.88)	(0.56, 0.79)	(0.51, 0.76)	(0.62, 0.82)
Big	0.67	0.64	0.62	0.63	0.68
	(0.55, 0.75)	(0.51, 0.71)	(0.54, 0.74)	(0.52, 0.76)	(0.69, 0.79)

B_{cf}	Growth	2	3	4	Value
Small	0.63	0.57	0.49	0.46	0.46
	(0.48, 0.86)	(0.45, 0.79)	(0.38, 0.71)	(0.37, 066)	(0.37, 0.63)
2	0.66	0.54	0.50	0.47	0.52
	(0.52, 0.83)	(0.42, 0.75)	(0.39, 0.67)	(0.40, 0.69)	(0.42, 0.67)
3	0.63	0.54	0.48	0.45	0.50
	(0.51, 0.80)	(0.46, 0.71)	(0.40, 0.68)	(0.38, 0.62)	(0.41, 0.61)
4	0.59	0.52	0.47	0.46	0.51
	(0.45, 0.81)	(0.40, 0.76)	(0.39, 0.67)	(0.38, 0.59)	(0.39, 0.63)
Value	0.48	0.46	0.42	0.43	0.47
	(0.41, 0.64)	(0.37, 0.66)	(0.35, 0.60)	(0.36, 0.58)	(0.40, 0.59)

Note: Estimated discount rate betas B_{dr} and cash-flow betas B_{cf} via the ICAPM-MS for value-size sorted portfolios for the modern sample, ranging from 03-1963 until 12-2022. The value in parentheses denote the 90% confidence intervals. The confidence intervals are constructed by by simulating 10.000 draws of the MSVAR parameters assuming a multivariate normal distribution for the estimates. Each iteration, the betas are re-estimated.

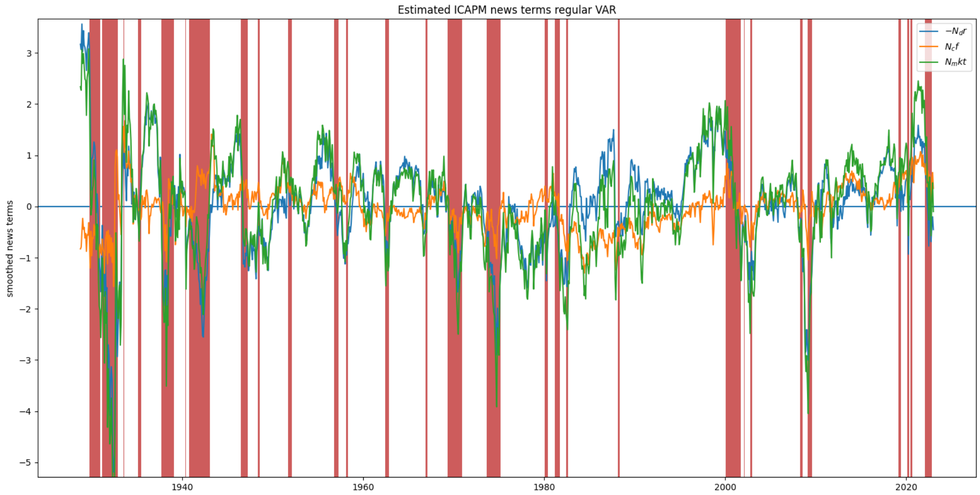

Figure 1: Estimated news terms of the VAR model and the MSVAR model, in percentage points. -N_{dr}, N_{cf} and N_{mkt} correspond to the (negative) discount rate, cash-flow and market news terms estimated by the VAR model. Similarly, -$N_{dr,ms}$, $N_{cf,ms}$ and $N_{mkt,ms}$ correspond to the (negative) discount rate, cash-flow and market news terms estimated by the MSVAR model. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5

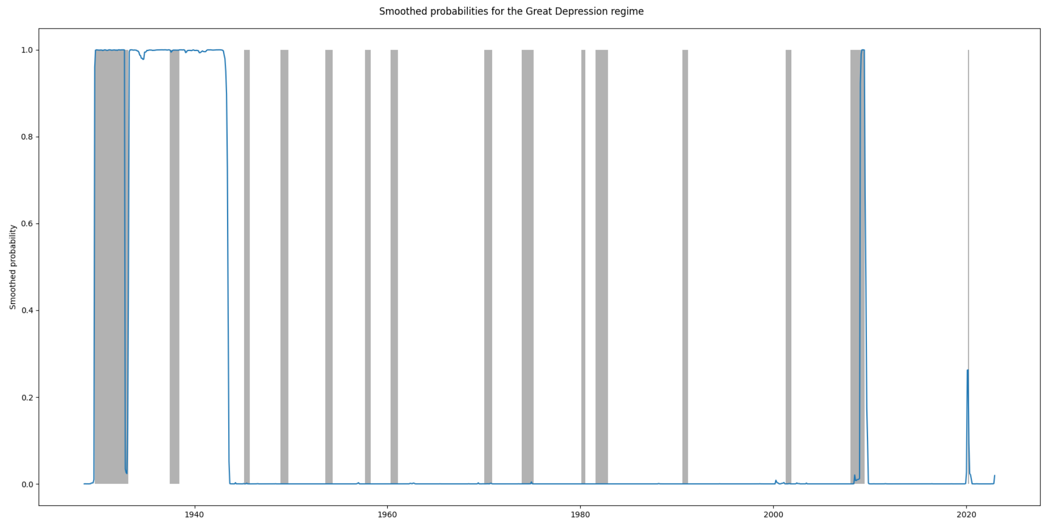

Figure 2: Smoothed probabilities of the Great Depression regime estimated via an MSVAR model assuming two level regimes and two variance regimes. The state variables are the excess market returns, small value spread, price to earning ratio and the dividend to price ratio, with monthly data from 07-1928 until 12-2022 (1334 observations). The shaded areas correspond to NBER recessions.

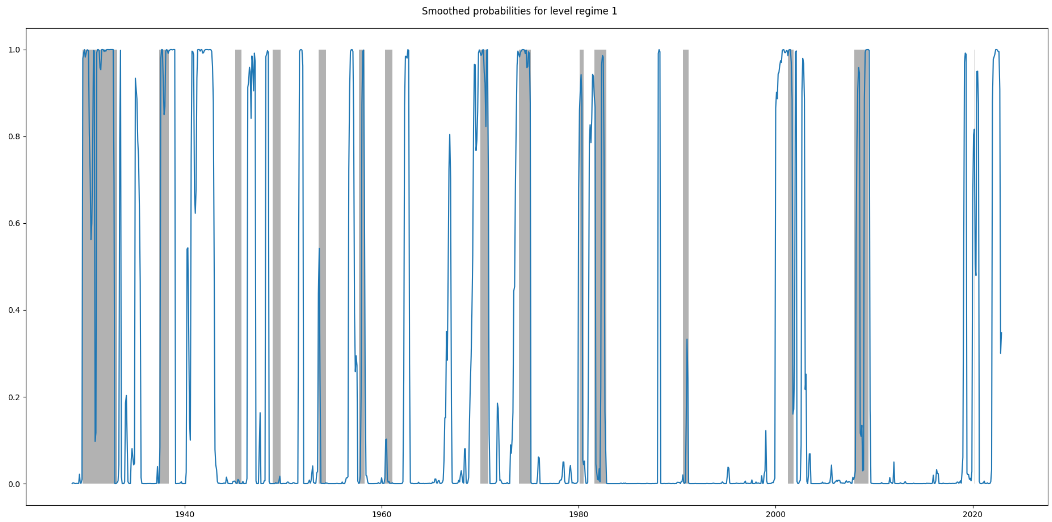

Figure 3: Smoothed probabilities of the crisis level regime estimated via an MSVAR model assuming two level regimes and two variance regimes, that are independent. The state variables are the excess market returns, term yield spread, small value spread, price to earning ratio and the dividend to price ratio, with monthly data from 07-1928 until 12-2022 (1334 observations). The shaded areas correspond to NBER recessions.

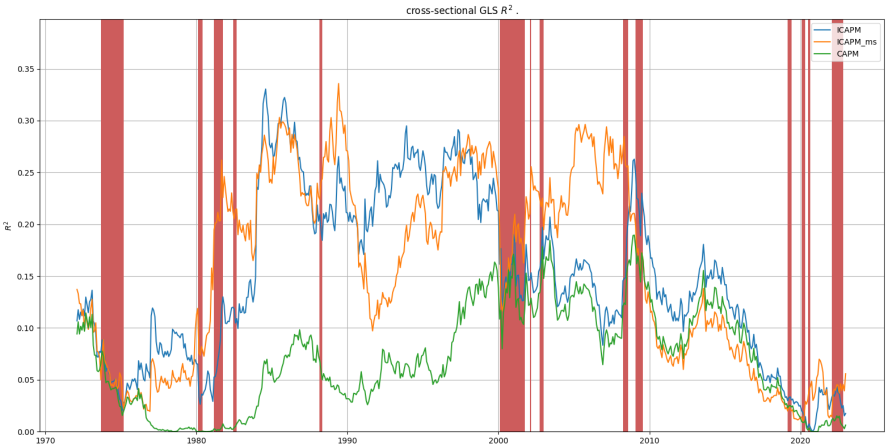

Figure 4: R^2 of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 25-year moving window from 01-1967 until 12-2022. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

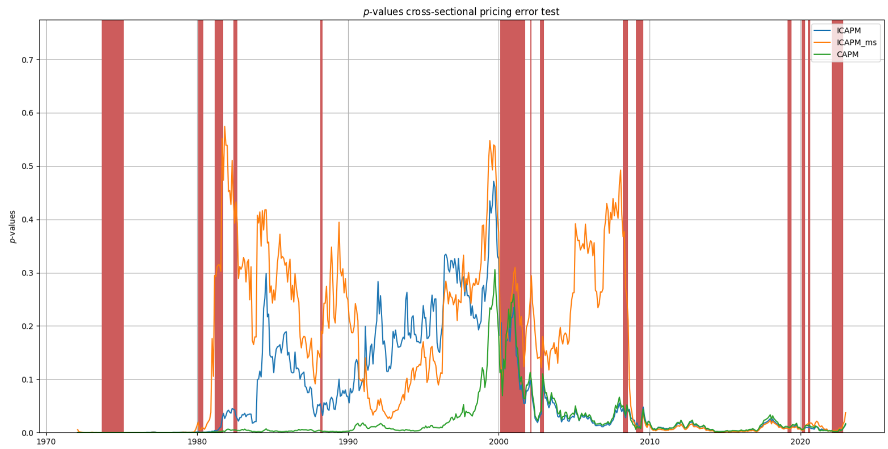

Figure 5: p-values of the asset pricing test of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 25-year moving window from 01-1967 until 12-2022-. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

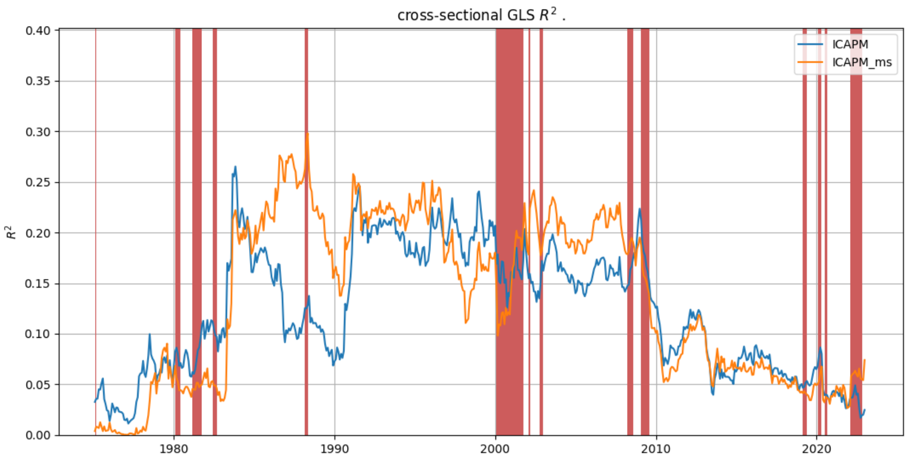

Figure 6: R^2 of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 45-year moving window from 1975-01 until 2022-12. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

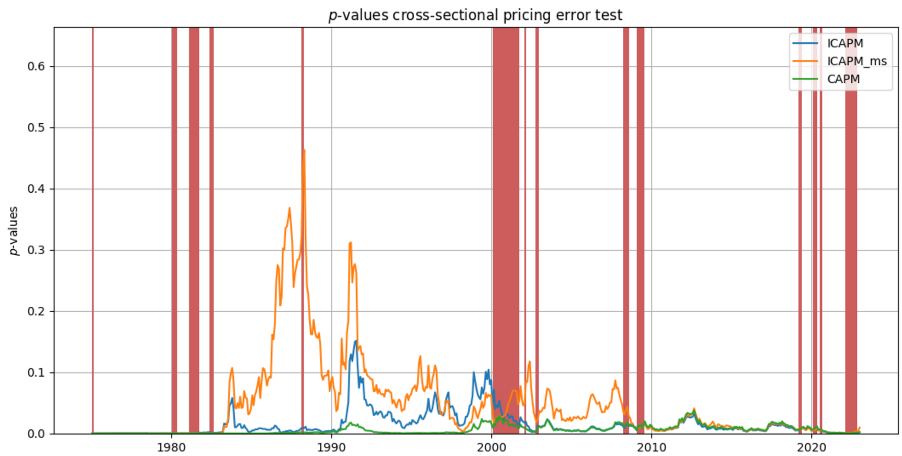

Figure 7: p-values of the asset pricing test of the second pass cross-sectional GLS regressions on 24 size-value sorted portfolios, using a 45-year moving window from 1975-01 until 2022-12. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.

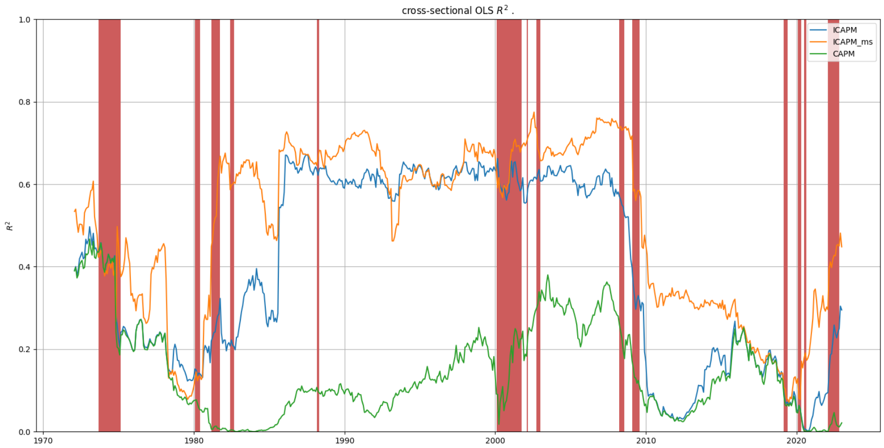

Figure 8: R^2 of the second pass cross-sectional OLS regressions on 24 size-value sorted portfolios, using a 35-year moving window from 1967-01 until 2022-12. The shaded areas represent the periods that the recession regimes prevail, during these periods the smoothed probabilities of the recession regime exceed 0.5.