Regime-switching factor models with applications to portfolio selection and demand estimation

Brian Zhu*

*Department of Economics, Yale University

6 April 2022

Abstract

I consider regime-switching factor models with the aim of capturing cyclical or latent-state-driven patterns in the market when explaining asset returns. Using the Baum-Welch algorithm to estimate model parameters, I derive new closed-form updates for non-Markov chain parameters that resemble weighted linear regressions. I fit these models on recent cryptocurrency data with market returns, size, and momentum as factors; results suggest that a moderate number of regimes are BIC-optimal for these factors. I provide two applications of these models: first, to select portfolios and record their out-of-sample returns; second, to simulate and estimate a cryptocurrency demand model.

1 Introduction

A regime-switching model associates observed data to latent variables representing the regime (e.g. a bull or bear market) that the observed data comes from. The latent regime variable may determine some or all parameters of a model for the data, and is typically modelled by a Markov chain. Thus, regime-switching models are also known as Markov-switching models and are cases of hidden Markov models. In financial economics, regime-switching models have been prized for their ability to capture cyclical patterns of financial markets. The first appearance of a regime-switching model in financial economics was in Hamilton (1989), where a regime-switching autoregressive model was

I firstly thank my advisor Professor Xiaohong Chen for her guidance and mentorship. I would also like to thank Professor Eduardo Dávila and the Tobin RA program for introducing me to research in financial economics. Finally, I thank my friends and family for their support. All errors are my own; data and code are available upon request.

used to detect business cycle breakpoints. Another application of the regime-switching framework is in regime-switching conditional heteroskedasticity models, first proposed by Hamilton and Sumsel (1994), where the path dependence of the conditional variance on the regime raises a tractability problem and thus has been the subject of further research (see Augustyniak (2014) for expectation-maximization algorithms or Billio et al. (2016) for Gibbs sampling methods, for example).

A factor model relates N variables, typically macroeconomic indicators or asset returns, to M factors by fitting a loading matrix, a diagonal idiosyncratic noise (or risk) matrix, and sometimes an intercept vector (called the alpha or excess return when modelling asset returns). When $N >> M$, factor models can be seen as a form of dimension reduction. Some factor models for asset returns are the capital asset pricing model (CAPM), first published in Sharpe (1964), with market returns as the sole factor, and the Fama-French three-factor model of Fama and French (1993), with market returns, size (market capitalization), and value (book-to-market ratio) as factors. The first appearance of a regime-switching factor model was in Kim and Yoo (1995), where a macroeconomic time series y_t was modeled as $y_t = \Lambda f_t + u_t$ where the factors f_t and unobserved components u_t were further modeled with regime-switching vector autoregressive models. Since then, regime-switching models have appeared in financial contexts, with between two (Nystrup et al. (2018)) and four (Bae et al. (2014)) regimes identified as suitable for best explaining stock returns. Costa and Kwon (2019) and Costa and Kwon (2020) proposed models $r_t = \alpha + \sum_s B_s f_t \mathbb{I}_s + \varepsilon_t$ and $r_t = \sum_s (\alpha_s + B_s f_t + \varepsilon_{s,t})\mathbb{I}_s$, respectively ($s$ denotes a state-dependent parameter or indicator function, ε_t is Gaussian noise, and $\varepsilon_{s,t}$ is state-dependent Gaussian noise), for the purpose of improving estimates of the mean vector and covariance matrix of asset returns for mean-variance portfolio optimization.

In this paper, I consider four regime-switching factor models based on Costa and Kwon (2019) and Costa and Kwon (2020). To estimate the parameters of these models, I use the Baum-Welch algorithm, an expectation-maximization algorithm for hidden Markov models. The forward-backward algorithm is used for the expectation step. For the maximization step, Baum et al. (1970) provide closed-form updates for the Markov chain parameters (initial distribution and transition matrix). The updates for the non-Markov chain parameters are dependent on the model specification. For each type of regime-switching factor model considered, I derive closed-form updates for the non-Markov chain parameters, which resemble weighted linear regressions. In deriving the closed-form updates, it is crucial to assume that the error term $\varepsilon_{s,t}$ is Gaussian conditional on the latent state s if idiosyncratic risk is modeled as state-dependent, and that ε_t is Gaussian if idiosyncratic risk is not modeled as state-dependent. To my knowledge, this is the first time that such closed-form updates have been explicitly characterized for these types of regime-switching factor models. The closed forms are useful particularly when the number of parameters is large, as matrix computations are much faster than numerical optimization procedures. It is also notable how regression-like procedures can still be used to fit factor models when a regime-switching component is added.

I fit these models on recent cryptocurrency data. Simply speaking, cryptocurrencies are digital currencies, i.e. binary data designed to serve as a medium of exchange and/or storage of value. The "crypto" part refers to the cryptographic blockchain technology that is used to record transactions. Notably, certain cryptocurrencies use proof-of-work, a decentralized form of control via networks of computers "mining" cryptocurrencies, i.e. verifying transactions, as opposed to central bank digital currencies or traditional currencies under the authority of a central bank. While Liu and Tsyvinski (2021) suggest that cryptocurrency returns and traditional currency returns behave differently and cast doubt that cryptocurrencies fulfill the two aforementioned purposes of a currency, the cryptocurrency market has expanded substantially in value and notoriety, especially over the past two years. The total market capitalization of cryptocurrencies has grown from circa $200 billion in November 2019 to over $2500 billion as of November 2021, and the total number of cryptocurrencies has nearly tripled from 2817 to 7557 over the same time period.[1,2] Social media and celebrities are among the causes of sudden periods of high volatility in cryptocurrency prices. These peculiarities suggest that regime-switching models are well-suited for cryptocurrency price data.

On explaining cryptocurrency returns, Urquhart (2016) finds that Bitcoin returns are predictable, with interesting examples of good predictors including Google Trends data (Urquhart (2018)) and Twitter activity (Shen et al. (2019)). This has motivated research on factor models to find risk factors that may explain cryptocurrency returns. Hubrich (2017) proposes a four-factor model of market returns, value (transaction volume to market capitalization ratio), momentum (prior returns), and carry (discrepancies between spot and futures markets), and Shen et al. (2020) examines a three-factor model of market returns, size, and reversals (prior returns). Given the recent growth and volatility of the cryptocurrency market, a factor model analysis with more recent data and an introduction of regime-switching framework with the goal of learning said volatility is relevant and is explored in Section 3 of this paper. Following the literature, I use market returns, size, and momentum as factors. I show that these are suitable factors for the data before fitting and analyzing regime-switching factor models.

[1] https://www.statista.com/statistics/730876/cryptocurrency-maket-value/
[2] https://www.statista.com/statistics/863917/number-crypto-coins-tokens/

I provide two interesting applications of these regime-switching factor models: portfolio selection and demand estimation. In portfolio selection, uncertainty the estimated mean vector and covariance matrix of asset returns has been a major issue. To mitigate this issue, for example, Goldfarb and Iyengar (2004) include uncertainty sets for the estimated parameters in the portfolio optimization problems, and Michaud and Michaud (2008) suggest resampling methods. Regime-switching factor models can also help resolve this issue, as well-chosen factors have the ability to quantifiably explain systematic risk in an asset market, and a regime-switching component further refines the estimated systemic and idiosyncratic risk of asset returns. Costa and Kwon (2019) and Costa and Kwon (2020) show that portfolios informed by regime-switching factor models outperform competing portfolios in out-of-sample stock returns. In Section 4, I estimate mean vectors and covariance matrices of asset returns via regime-switching factor models to construct portfolios for an investor with mean-variance utility and short-selling constraints. Numerical experiments show that portfolios constructed this way outperform those constructed using other common estimates of the asset return distribution and common trading strategies under the same rebalancing and short-selling restrictions.

Beliefs about economic conditions are important in shaping investors' portfolios (Giglio et al. (2021)). Benetton and Compiani (2020) include investor beliefs about cryptocurrencies, measured by survey responses, as covariates in the heterogeneous agent model for asset pricing of Koijen and Yogo (2019) to estimate the effect on cryptocurrency demand. In Section 5, I use regime-switching factor models to simulate and estimate the model, focusing on how investors' traits and belief biases, i.e. the bandwagon effect, bias towards volatile states, and risk aversion, affect cryptocurrency demand. While Benetton and Compiani (2020) use holdings data, I simulate holdings data by estimating latent variable distributions from the regime-switching factor model, tilting the distribution by belief bias parameters (randomly drawn for each simulated investor), and constructing portfolios for each simulated investor under mean-variance utility weighted by risk aversion and short-selling constraints. The rationale behind this approach is that the estimated latent variable distributions are a proxy for the true market state distribution, and that investors are rational given their perception of the market state distribution, which I model with the estimated latent variable distribution tilted by their belief bias parameters. Under these assumptions, heterogeneity and pricing anomalies result from differing traits and belief bias parameters. With the simulated data set as the response, and the belief bias parameters as covariates in the model of Koijen and Yogo (2019), coefficients of those parameters and control variables can be estimated. Results show that the coefficients of the bandwagon effect and risk aversion parameters are negative with statistical significance.

The remainder of this paper is organized as follows: Section 2 describes the regime-switching factor models and methods of estimation and inference for those models; Section 3 describes the cryptocurrency data and factor construction process, and analyzes the regime-switching factor models fitted using the data; Sections 4 and 5 cover the applications to portfolio selection and demand estimation applications, respectively, and Section 6 concludes.

2 Regime-Switching Factor Models

2.1 Model Description

I would like to model an asset market of N assets with S possible latent states and M factors using T observation periods. Let $\pi_{t,i}$ and r_t^f be the observed price of coin i and risk-free rate, respectively, at time t. The excess return on asset i at time t is

$$r_{t,i} = 100\left(\frac{\pi_{t,i}}{\pi_{t-1,i}} - 1\right) - r_t^f$$

Let $r_t = (r_{t,1}, \ldots, r_{t,1})^\top$. Let $f_t \in \mathbb{R}^M$ be the observed factor vector and z_t be the latent realization of a Markov chain $\{Z_t\}$ on $\{1, \ldots, S\}$, respectively, at time t. I define four regime-switching factor models, based on those of Costa and Kwon (2019) and Costa and Kwon (2020), as follows:

- The *regime-switching factor model (RSFM)* is

  $$r_t = \sum_{s=1}^{S} (\alpha_s + B_s f_t + \varepsilon_{s,t}) \cdot \mathbb{I}[z_t = s]$$

  with parameters $\alpha_s \in \mathbb{R}^N$ (state-dependent excess returns), $B_s \in \mathbb{R}^{N\times M}$ (state-dependent factor loadings), and $\sigma_s \in \mathbb{R}^N$ (state-dependent idiosyncratic risk such that $\varepsilon_{s,t} \sim \mathcal{N}(0, \operatorname{diag}\sigma_s^2)$ i.i.d.) for each $s \in \{1, \ldots, S\}$, $p \in \mathbb{R}^S$ (initial distribution), and $P \in \mathbb{R}^{S\times S}$ (transition matrix).

- The *alpha-restricted regime-switching factor model (α-RSFM)* is

  $$r_t = \alpha + \sum_{s=1}^{S} (B_s f_t + \varepsilon_{s,t}) \cdot \mathbb{I}[z_t = s]$$

  with parameters $B_s \in \mathbb{R}^{N\times M}$ (state-dependent factor loadings), and $\sigma_s \in \mathbb{R}^n$ (state-dependent idiosyncratic risk such that $\varepsilon_{s,t} \sim \mathcal{N}(0, \operatorname{diag}\sigma_s^2)$ i.i.d.) for each $s \in \{1, \ldots, S\}$, $\alpha \in \mathbb{R}^N$ (excess returns), $p \in \mathbb{R}^S$ (initial distribution), and $P \in \mathbb{R}^{S\times S}$ (transition matrix).

- The *sigma-restricted regime-switching factor model (σ-RSFM)* is

$$r_t = \sum_{s=1}^{S}((\alpha_s + B_s f_t) \cdot \mathbb{I}[z_t = s]) + \varepsilon_t$$

  with parameters $\alpha_s \in \mathbb{R}^N$ (state-dependent excess returns), $B_s \in \mathbb{R}^{N \times M}$ (state-dependent factor loadings) for each $s \in \{1, \ldots, S\}$, $\sigma \in \mathbb{R}^N$ (idiosyncratic risk such that $\varepsilon_t \sim \mathcal{N}(0, \operatorname{diag} \sigma^2)$ i.i.d.), $p \in \mathbb{R}^S$ (initial distribution), and $P \in \mathbb{R}^{S \times S}$ (transition matrix).

- The *alpha- and sigma-restricted regime-switching factor model ($\alpha\sigma$-RSFM)* is

$$r_t = \alpha + \sum_{s=1}^{S}(B_s f_t \cdot \mathbb{I}[z_t = s]) + \varepsilon_t$$

  with parameters $B_s \in \mathbb{R}^{N \times M}$ (state-dependent factor loadings) for each $s \in \{1, \ldots, S\}$, $\alpha \in \mathbb{R}^N$ (excess returns), $\sigma \in \mathbb{R}^N$ (idiosyncratic risk such that $\varepsilon_t \sim \mathcal{N}(0, \operatorname{diag} \sigma^2)$ i.i.d.), $p \in \mathbb{R}^S$ (initial distribution), and $P \in \mathbb{R}^{S \times S}$ (transition matrix).

2.2 Model Estimation

I would like to estimate the parameters of the regime-switching factor models with the expectation-maximization (EM) algorithm, an iterative algorithm where each iteration consists of an expectation (E) step and a maximization (M) step. Let θ be the parameter vector, and define $z = (z_1, \ldots, z_T)$, $r = (r_1, \ldots, r_T)$, and $f = (f_1, \ldots, f_T)$. Given the initial or a previous θ_0, the E-step computes the distribution $(z|r, f, \theta_0)$, and the M-step maximizes the evidence lower bound $\mathcal{Q}(\theta|\theta_0)$ given by

$$\begin{aligned}
\mathcal{Q}(\theta|\theta_0) = \mathbb{E}_{z|r,f,\theta_0}[\log \mathbb{P}[r, z|f, \theta]] &= \mathbb{E}_{z|r,f,\theta_0}\left[\sum_{t=1}^{T} \log \mathbb{P}[r_t, z_t|f, \theta]\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{z|r,f,\theta_0}[\log(\mathbb{P}[z_t|f, \theta] \cdot \mathbb{P}[r_t|z_t, f, \theta])] \\
&= \sum_{t=1}^{T}\sum_{s=1}^{S} \mathbb{P}[z_t = s|r, f, \theta_0](\log \mathbb{P}[z_t = s|f, \theta] + \log \mathbb{P}[r_t|z_t = s, f, \theta])
\end{aligned}$$

2.2.1 Expectation Step

Let $q_t(s) = \mathbb{P}[z_t = s|r, f, \theta_0]$ for time t and state s. Note that computing $q_t(s)$ for all $t \in \{1, \ldots, T\}$ and $s \in \{1, \ldots, S\}$ gives the distribution of $(z|r, f, \theta_0)$, thus completing the E-step. I compute the

$q_t(s)$ with the Baum-Welch (BW) algorithm: let $\zeta_{s,t} = \mathbb{P}[r_1, \dots, r_t, z_t = s | f, \theta_0]$ be the forward density and $\xi_{s,t} = \mathbb{P}[r_{t+1}, \dots, r_T | z_t = s, f, \theta_0]$ be the backward density for time t and state s; then

$$\zeta_{s,1} = p_0(s) \cdot \mathbb{P}[r_t | z_t = s, f, \theta_0]$$

$$\zeta_{s,t} = \sum_{s'=1}^{S} \zeta_{s',t-1} \cdot \mathbb{P}[r_t | z_t = s', f, \theta_0] \cdot P_0(s', s) \text{ for } t = 2, \dots, T$$

$$\xi_{s,T} = 1$$

$$\xi_{s,t} = \sum_{s'=1}^{S} \xi_{s,t+1} \cdot P_0(s, s') \cdot \mathbb{P}[r_{t+1} | z_{t+1} = s', f, \theta_0] \text{ for } t = T-1, \dots, 1$$

Let $q_t(s, s') = \mathbb{P}[z_t = s, z_{t+1} = s' | r, f, \theta_0]$. Then

$$q_t(s) = \frac{\zeta_{s,t}\xi_{s,t}}{\sum_{s'=1}^{S} \zeta_{s',t}\xi_{s',t}}$$

$$q_t(s, s') = \frac{\zeta_{s,t} \cdot P_0(s, s') \cdot \xi_{s',t} \cdot \mathbb{P}[r_{t+1} | z_{t+1} = s', f, \theta_0]}{\sum_{\varsigma=1}^{S} \sum_{\varsigma'=1}^{S} \zeta_{\varsigma,t} \cdot P_0(\varsigma, \varsigma') \cdot \xi_{\varsigma',t} \cdot \mathbb{P}[r_{t+1} | z_{t+1} = \varsigma', f, \theta_0]}$$

2.2.2 Maximization Step

Given $q_t(s)$ for all $t \in \{1, \dots, T\}$ and $s \in \{1, \dots, S\}$, I write $\mathcal{Q}(\theta|\theta_0)$ as

$$\begin{aligned}
\mathcal{Q}(\theta|\theta_0) &= \sum_{t=1}^{T} \sum_{s=1}^{S} q_t(s) \log \mathbb{P}[z_t = s | f, \theta] + \sum_{t=1}^{T} \sum_{s=1}^{S} q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta] \\
&= \sum_{t=1}^{T} \sum_{s=1}^{S} q_t(s) \log \big(P^t p\big)_s + \sum_{t=1}^{T} \sum_{s=1}^{S} q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta]
\end{aligned}$$

noting that knowing f alone gives no further information about z_t, so $\mathcal{Q}(\theta|\theta_0)$ is separable w.r.t. the Markov chain and factor model parameters. Then solving the two induced sub-problems completes the M-step. The BW algorithm gives the Markov chain parameter updates as

$$p^\star(s) = q_1(s)$$

$$P^\star(s, s') = \frac{\sum_{t=1}^{T-1} q_t(s, s')}{\sum_{t=1}^{T-1} q_t(s)}$$

I now give closed-form expressions for the factor model parameters, based on the type of model.

- Regime-switching factor model (RSFM): define

$$
\begin{aligned}
\tilde{q}_s &= ((q_1(s))^{1/2}, \ldots, (q_T(s))^{1/2})^\top \\
\tilde{r}_{s,i} &= ((q_1(s))^{1/2} r_{1,i}, \ldots, (q_T(s))^{1/2} r_{T,i})^\top \\
\Phi_s &= \begin{bmatrix} (q_1(s))^{1/2} & (q_1(s))^{1/2} f_{1,1} & \cdots & (q_1(s))^{1/2} f_{1,M} \\ \vdots & \vdots & \ddots & \vdots \\ (q_T(s))^{1/2} & (q_T(s))^{1/2} f_{T,1} & \cdots & (q_T(s))^{1/2} f_{T,M} \end{bmatrix}
\end{aligned}
$$

Let $B_{s,i}$ be the i-th row of B_s. Then

$$
\begin{aligned}
(\alpha^\star_{s,i}, B^\star_{s,i}) &= (\Phi_s^\top \Phi_s)^{-1} \Phi_s^\top \tilde{r}_{s,i} \\
\sigma^\star_{s,i} &= \frac{\|\tilde{r}_{s,i} - \Phi_s(\alpha^\star_{s,i}, B^\star_{s,i})\|_2}{\|\tilde{q}_s\|_2}
\end{aligned}
$$

- Alpha-restricted regime-switching factor model (α-RSFM): define

$$
\begin{aligned}
q_s &= (q_1(s), \ldots, q_T(s))^\top \\
r_{s,i} &= (q_1(s) \cdot r_{1,i}, \ldots, q_T(s) \cdot r_{T,1})^\top \\
F &= \begin{bmatrix} f_{1,1} & \cdots & f_{1,M} \\ \vdots & \ddots & \vdots \\ f_{T,1} & \cdots & f_{T,M} \end{bmatrix} \\
\tilde{F}_s &= \begin{bmatrix} (q_1(s))^{1/2} f_{1,1} & \cdots & (q_1(s))^{1/2} f_{1,M} \\ \vdots & \ddots & \vdots \\ (q_1(s))^{1/2} f_{T,1} & \cdots & (q_T(s))^{1/2} f_{T,M} \end{bmatrix}
\end{aligned}
$$

Then

$$
\begin{aligned}
\alpha^\star_i &= \frac{\sum_{t=1}^T r_{t,i} - \sum_{s=1}^S (F^\top r_{s,i})^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)}{T - \sum_{s=1}^S (F^\top q_s)^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)} \\
B^\star_{s,i} &= (\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha^\star_i q_s) \\
\sigma^\star_{s,i} &= \left(\frac{\sum_{t=1}^T q_s(t)(r_{t,i} - \alpha^\star_i - B^{\star\top}_{s,i} f_t)^2}{\sum_{t=1}^T q_t(s)} \right)^{1/2}
\end{aligned}
$$

- Sigma-restricted regime-switching factor model (σ-RSFM):

$$
\begin{aligned}
(\alpha^\star_{s,i}, B^\star_{s,i}) &= (\Phi_s^\top \Phi_s)^{-1} \Phi_s^\top \tilde{r}_{s,i} \\
\sigma^\star_i &= \left(\frac{\sum_{s=1}^S \|\tilde{r}_{s,i} - \Phi_s(\alpha^\star_{s,i}, B^\star_{s,i})\|_2}{T} \right)^{1/2}
\end{aligned}
$$

- Alpha- and sigma-restricted regime-switching factor model ($\alpha\sigma$-RSFM):

$$
\begin{aligned}
\alpha^\star_i &= \frac{\sum_{t=1}^T r_{t,i} - \sum_{s=1}^S (F^\top r_{s,i})^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)}{T - \sum_{s=1}^S (F^\top q_s)^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)} \\
B^\star_{s,i} &= (\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha^\star_i q_s) \\
\sigma^\star_i &= \left(\frac{\sum_{s=1}^S \sum_{t=1}^T q_s(t)(r_{t,i} - \alpha^\star_i - B^{\star\top}_{s,i} f_t)^2}{T} \right)^{1/2}
\end{aligned}
$$

Derivations and practical considerations for computing these closed-form expressions are in the appendix.

2.3 Model Inference

I summarize Kuan (2002) which describes inference on regime-switching models using prediction, filtering, and smoothing probabilities. Let $\mathcal{R}_t = (r_1, \ldots, r_t)$ and $r = (r_1, \ldots, r_T)$. The prediction probabilities $\mathbb{P}[z_t = s | \mathcal{R}_{t-1}, \theta]$ predict the latent variable distributions given past information and satisfy

$$
\mathbb{P}[z_t = s | \mathcal{R}_{t-1}, \theta] = \sum_{s'=1}^S \mathbb{P}[z_{t-1} = s' | \mathcal{R}_{t-1}, \theta] \cdot P(s', s)
$$

where $P(s', s)$ is the (s', s)-th entry of P. The filtering probabilities $\mathbb{P}[z_t = s | \mathcal{R}_t, \theta]$ give the latent variable distributions given past and current information, and satisfy

$$
\mathbb{P}[z_t = s | \mathcal{R}_t, \theta] = \frac{\mathbb{P}[z_t = s | \mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t | z_t = s, \mathcal{R}_{t-1}, \theta]}{\sum_{s'=1}^S \mathbb{P}[z_t = s' | \mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t | z_t = s', \mathcal{R}_{t-1}, \theta]}
$$

where $\mathbb{P}[r_t | z_t = s, \mathcal{R}_{t-1}, \theta]$ is the conditional density of r_t given $z_t = s$, $\mathbb{R}_{t=1}$, and θ. Given p and P, the prediction and filtering probabilities can be recursively computed starting from $t = 1$ and going forward as follows:

$$
\begin{aligned}
\mathbb{P}[z_1 = s|\mathcal{R}_0, \theta] &= p(s) \\
\mathbb{P}[z_1 = s|\mathcal{R}_1, \theta] &= \frac{\mathbb{P}[z_1 = s|\mathcal{R}_0, \theta] \cdot \mathbb{P}[r_1|z_1 = s, \mathcal{R}_0, \theta]}{\sum_{s'=1}^{S} \mathbb{P}[z_1 = s'|\mathcal{R}_0, \theta] \cdot \mathbb{P}[r_1|z_1 = s', \mathcal{R}_0, \theta]} \\
\mathbb{P}[z_2 = s|\mathcal{R}_1, \theta] &= \sum_{s'=1}^{S} \mathbb{P}[z_1 = s'|\mathcal{R}_1, \theta] \cdot P(s', s) \\
\mathbb{P}[z_2 = s|\mathcal{R}_2, \theta] &= \frac{\mathbb{P}[z_2 = s|\mathcal{R}_1, \theta] \cdot \mathbb{P}[r_2|z_2 = s, \mathcal{R}_1, \theta]}{\sum_{s'=1}^{S} \mathbb{P}[z_2 = s'|\mathcal{R}_1, \theta] \cdot \mathbb{P}[r_2|z_1 = s', \mathcal{R}_1, \theta]} \\
&\vdots \\
\mathbb{P}[z_T = s|\mathcal{R}_{T-1}, \theta] &= \sum_{s'=1}^{S} \mathbb{P}[z_{T-1} = s'|\mathcal{R}_{T-1}, \theta] \cdot P(s', s) \\
\mathbb{P}[z_T = s|\mathcal{R}_{T-1}, \theta] &= \frac{\mathbb{P}[z_T = s|\mathcal{R}_{T-1}, \theta] \cdot \mathbb{P}[r_T|z_T = s, \mathcal{R}_{T-1}, \theta]}{\sum_{s'=1}^{S} \mathbb{P}[z_T = s'|\mathcal{R}_{T-1}, \theta] \cdot \mathbb{P}[r_T|z_T = s', \mathcal{R}_{T-1}, \theta]} \\
\mathbb{P}[z_{T+1} = s|\mathcal{R}_T, \theta] &= \sum_{s'=1}^{S} \mathbb{P}[z_T = s|\mathcal{R}_T, \theta] \cdot P(s', s)
\end{aligned}
$$

where $p(s)$ is the s-th entry of p. The smoothing probabilities $\mathbb{P}[z_t = s|r, \theta]$ give the latent variable distributions given all information and, per Kim (1994), approximately satisfy

$$
\mathbb{P}[z_t = s|r, \theta] \approx \mathbb{P}[z_t = s|\mathcal{R}_t, \theta] \sum_{s'=1}^{S} \frac{P(s, s') \cdot \mathbb{P}[z_{t+1} = s'|r, \theta]}{\mathbb{P}[z_{t+1} = s'|\mathcal{R}_t, \theta]}
$$

Given the prediction and filtering probabilities, the smoothing probabilities can be recursively computed starting from $t = T$ and going backwards as follows:

$$
\begin{aligned}
\mathbb{P}[z_T = s|r, \theta] &= \mathbb{P}[z_T = s|\mathcal{R}_T, \theta] \sum_{s'=1}^{S} \frac{P(s, s') \cdot \mathbb{P}[z_{T+1} = s'|r, \theta]}{\mathbb{P}[z_{T+1} = s'|\mathcal{R}_T, \theta]} \\
\mathbb{P}[z_{T-1} = s|r, \theta] &= \mathbb{P}[z_{T-1} = s|\mathcal{R}_{T-1}, \theta] \sum_{s'=1}^{S} \frac{P(s, s') \cdot \mathbb{P}[z_T = s'|r, \theta]}{\mathbb{P}[z_T = s'|\mathcal{R}_{T-1}, \theta]} \\
&\vdots \\
\mathbb{P}[z_1 = s|r, \theta] &= \mathbb{P}[z_1 = s|\mathcal{R}_1, \theta] \sum_{s'=1}^{S} \frac{P(s, s') \cdot \mathbb{P}[z_2 = s'|r, \theta]}{\mathbb{P}[z_2 = s'|\mathcal{R}_1, \theta]}
\end{aligned}
$$

Note that $\mathbb{P}[z_{T+1} = s'|r, \theta] = \mathbb{P}[z_{T+1} = s'|\mathcal{R}_T, \theta]$ in the first equation.

3 Cryptocurrency Data and Factors

3.1 Data Description

I collect cryptocurrency data, i.e. prices and market capitalizations, from CoinGecko and one-month treasury bill rates from the U.S. Treasury Department.[3,4] I use a sample period from 14 October 2020 to 13 December 2021 with daily frequency, yielding a total of 425 observations. The first 7 observations are used to calibrate momentum and the remaining observations are used to fit models, so $T = 418$. I select the 100 cryptocurrencies with the largest market capitalizations at the end of the sample period that have available data over the entire sample period, so $N = 100$. Since many cryptocurrencies have been created or have failed within the past two years, this selection process yields a basket of cryptocurrencies reasonably representative of the cryptocurrency market. Tables 1 and 2 display summary statistics in the aggregate and for specific cryptocurrencies.

Table 1: Aggregate Summary Statistics

Mean	StDv	Skew	Kurt	5th Pctile	95th Pctile
0.7861	8.5913	3.5856	78.8289	-10.2919	13.8014

Table 2: Summary Statistics for Selected Coins

Coin Name	Mean	StDv	Skew	Kurt	5th Pctile	95th Pctile
Bitcoin	0.3666	4.1150	0.1050	1.5513	-6.1461	7.6039
Ethereum	0.6681	5.3950	-0.0810	2.9632	-7.5025	8.9280
Dogecoin	1.9730	20.1518	11.8330	187.4399	-10.3864	18.7804

The aggregate summary statistics show that the distribution of cryptocurrency returns is right-skewed and very heavy-tailed. Examination of the summary statistics of specific coins show that there are coins with relatively nonvolatile, symmetric, light-tailed distributions such as Bitcoin and Ethereum, which are both considered "established" in the community, as well as coins with volatile, right-skewed, extremely heavy-tailed distributions such as Dogecoin, considered a "meme" coin by many in the community.

[3] https://www.coingecko.com/en/api
[4] https://home.treasury.gov/

3.2 Factor Construction

I consider a three-factor model of the market, size, and momentum, inspired by Hubrich (2017) and Shen et al. (2020). The market factor at time t, denoted by MKT_t, is the excess return on the market portfolio, a value-weighted portfolio of all cryptocurrencies, at time t. Formally, let $c_{t,i}$ be the market capitalization of coin i at time t, and let $c_t = c_{t,1} + \cdots + c_{t,N}$. Then

$$\mathrm{MKT}_t = \sum_{i=1}^{N} r_{i,t} \frac{c_{i,t}}{c_t}$$

The size and momentum factors are constructed similarly to the factors in Fama and French (1993). The size factor at time t, denoted by SMB_t, is the excess return on the small-minus-big portfolio, a value-weighted portfolio long in small (low market capitalization) stocks and short in big (high market capitalization) stocks, a time t. The momentum factor at time t, denoted by WML_t, is the excess return on the winner-minus-loser portfolio, a value-weighted portfolio long in winner (high average prior return) stocks and short in loser (low average prior return) stocks. Formally, I set a holding period of $H = 7$ (days). Before the start of the sample period and at the end of every holding period, I sort the coins by market capitalization at that time and by average prior return over the past H periods. Let B ("big") and S ("small") contain the top 20 and remaining 80 coins by market capitalization, respectively. Let W ("winner"), M ("middle"), and L ("loser") contain the top 30, middle 40, and bottom 30 coins by average prior returns. Let BW, BM, BL, SW, SM, and SL be value-weighted portfolios of the corresponding intersected coin sets. Denoting excess returns on the portfolios at time t with a t subscript, the size and momentum factors are, respectively,

$$\begin{aligned}
\mathrm{SMB}_t &= \frac{1}{3}(\mathrm{SW}_t + \mathrm{SM}_t + \mathrm{SL}_t) - \frac{1}{3}(\mathrm{BW}_t + \mathrm{BM}_t + \mathrm{BL}_t) \\
\mathrm{WML}_t &= \frac{1}{2}(\mathrm{BW}_t + \mathrm{SW}_t) - \frac{1}{2}(\mathrm{BL}_t + \mathrm{SL}_t)
\end{aligned}$$

I now provide empirical evidence justifying my choice of factors. First, at each period, I sort the cryptocurrencies into quintiles by market capitalization and one-week average prior return. Table 3 displays the average returns for each quintile, showing that on average, small coins have higher returns than large coins, and that winning coins have higher returns than losing coins. The paired t-tests show that period-for-period, differences in the returns of the extreme quintiles are significant, suggesting that size and momentum are a source of variability for cryptocurrency returns.

Table 3: Mean Quintile Portfolio Returns

	Q1	Q2	Q3	Q4	Q5	Q1-Q5 Paired t
Size	0.4828	0.6815	0.6922	0.8485	1.1886	4.5888***
Momentum	0.9047	0.6360	0.5493	0.4498	0.3288	2.4600**

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

Second, I fit the CAPM and the three-factor model of the market, size, and momentum on the data according to the factor construction process described above. Table 4 displays the absolute alphas (excess return), adjusted R-squared values (variability explained), and idiosyncratic volatilities averaged across all coins for both models, with paired t-test results for the differences.

Table 4: CAPM vs. Three-Factor Model

	$\lvert\alpha\rvert$	R^2_{adj}	σ
CAPM	0.3227	0.3771	6.6375
3FM	0.2920	0.4110	6.4343
Paired t	2.8602***	13.2450***	13.1082***

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

The results show that the absolute alphas and adjusted R-squared values are significantly higher, and that the idiosyncratic volatilities are significantly lower for the three-factor model compared to the CAPM. Thus, the added factors help explain the returns significantly better than the market factor alone, further demonstrating the suitability of these factors for the cryptocurrency market.

3.3 Results

3.3.1 Optimal Model and Number of Regimes

Two tasks of interest are to find the optimal type of model and the optimal umber of regimes for the data. To this end, I fit RSFMs, α-RSFMs, σ-RSFMs, and $\alpha\sigma$-RSFMs for $S \in \{1, 2, 3, 4, 5\}$ regimes using $K = 10$ random initializations and 50 iterations per initialization. I record the log-likelihood of the output θ_k of the k-th initialization, given by

$$\log L(\theta_k) = \log \mathbb{P}[r|\theta_k] = \sum_{t=1}^{T} \sum_{s=1}^{S} (\log \mathbb{P}[z_t = s|\theta_k] + \log \mathbb{P}[r_t|z_t = s, \theta_k])$$

and the Bayesian information criterion (BIC) of the log-likelihood-optimal output $\theta^\star$ given by

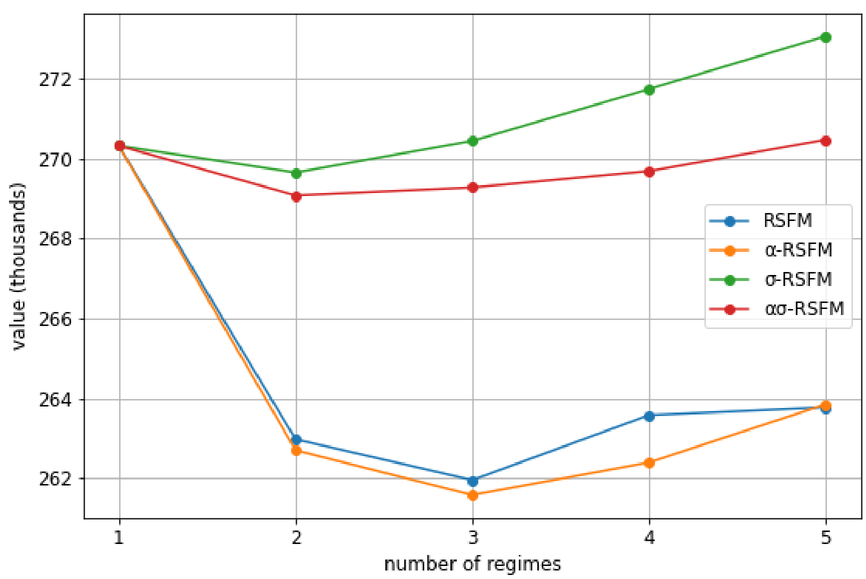
$$\text{BIC} = -2 \log L(\theta^\star) + \kappa(\theta^\star) \log T$$

where $\kappa(\theta^\star)$ is the number of free parameters of $\theta^\star$, given by

$$\kappa(\theta^\star) = \begin{cases} N(M+2)S + S^2 - 1 & = \text{ for the RSFM} \\ N(M+1)S + N + S^2 - 1 & = \text{ for the } \alpha\text{-RSFM and } \sigma\text{-RSFM} \\ NMS + 2N + S^2 - 1 & = \text{ for the } \alpha\sigma\text{-RSFM} \end{cases}$$

Figure 1 plots the BIC for each type of model over $S \in \{1, 2, 3, 4, 5\}$, suggesting that the α-RSFM with $S = 3$ regimes is BIC-optimal, with the three-regime RSFM not far behind. It is notable how restricting alpha generally improves the BIC for the RSFM and σ-RSFM.

Figure 1: Bayesian Information Criteria



However, as shown in Table 5, generalized likelihood ratio tests (GLRTs) of $H_0 : \alpha_1 = \cdots = \alpha_S$ for $S \in \{2, 3, 4, 5\}$ reject H_0, providing evidence in favor of the RSFM over the α-RSFM. A table of the BIC values is in the appendix. Thus different types of models may be optimal depending on the criteria used for evaluation.

Table 5: GLRTs of $H_0 : \alpha_1 = \cdots = \alpha_S$

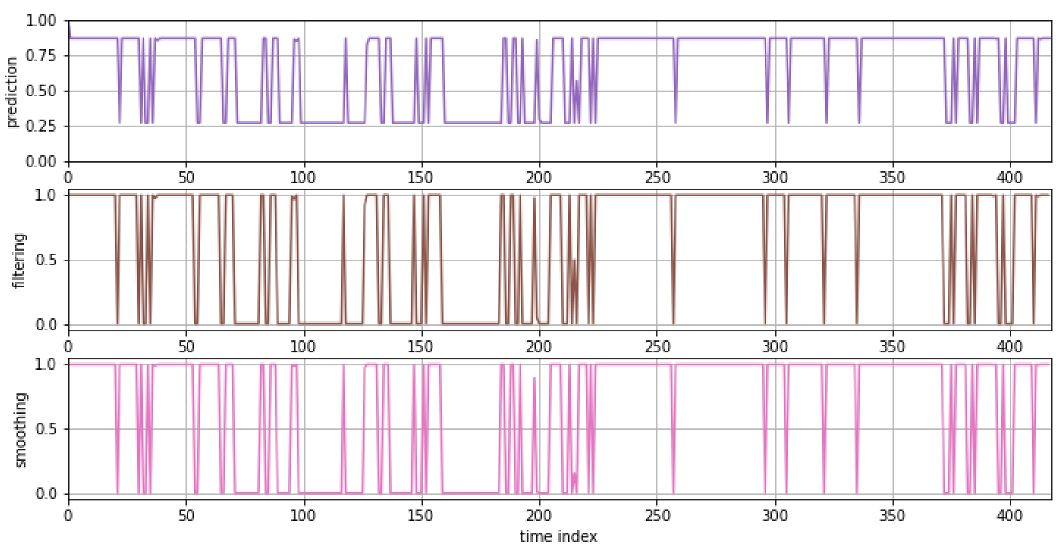
	Number of Regimes			
	2	3	4	5
Test Statistic (χ^2 degrees of freedom)	328.4981*** (100)	832.4128*** (200)	625.9918*** (300)	2489.9565*** (400)

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

3.3.2 Inference on the Latent Regime Variable

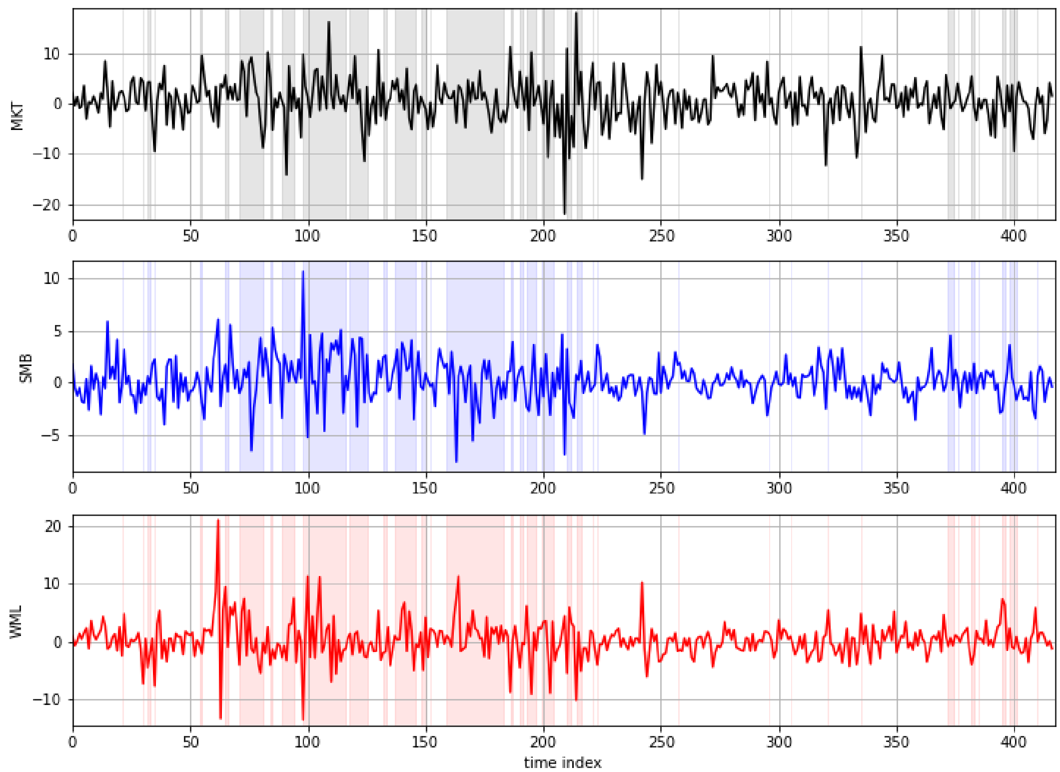
I demonstrate the inference procedures described in Section 2.3 on a two-regime RSFM with regimes indexed by $s \in \{1, 2\}$. Figure 2 plots the prediction, filtering, and smoothing probabilities for regime 1 across time.

Figure 2: Prediction, Filtering, and Smoothing Probabilities



The filtering and smoothing probabilities are very similar to each other and close to zero or one, suggesting that current information, once observed, plays a major role in determining the regime. Because of this, the prediction probabilities are very similar to corresponding entries of the estimated transition matrix. Figure 3 plots an equally-weighted average of the observed excess returns and time series of the factors with the likeliest regime, based on the smoothing probabilities, overlaid. A white background corresponds to regime 1 and a shaded background corresponds to regime 0.

Figure 3: Return and Factor Time Series with Likeliest Regime



It appears that regime 1 generally corresponds to periods where the factor time series are less volatile and regime 2 generally corresponds to periods where the factor time series are more volatile. Indeed, the average idiosyncratic risk in regime 1 is about 4.7189, while the idiosyncratic risk in regime 2 is about 8.4808. This suggests that regimes are differentiated primarily by the volatility of observed returns and factors. The figure suggests that the likeliest regime often switches rapidly. As there are no serial relations in the models considered, observations are assigned to regimes based on the values alone, not their position in time. Likeliest regime switches be more "smooth" and "connected" if serial relations are added between the observed returns and/or factors.

4 Portfolio Selection

4.1 Setting

Let μ and Σ be the mean vector and covariance matrix of joint returns on N assets. Suppose that an investor has mean-variance utility with risk aversion parameter $\gamma > 0$, i.e. the investor's utility from holding a portfolio specified by a weight vector w is given by the function

$$u(w) = \mu^\top w - \frac{\gamma}{2} w^\top \Sigma w$$

How to estimate μ and Σ has been a large topic in portfolio selection. With fitted regime-switching factor models, the estimate of μ and Σ at time t can be refined by conditioning on z_{t-1}. Let μ_f and Σ_f be the mean vector and covariance matrix, respectively, of the factors. Let $\hat{p}$ be the prediction probability vector for z_t at time $(t-1)$. Then (derivations are in the appendix),

$$\begin{aligned}
\mathbb{E}[r_t|z_{t-1}] &= \sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s \mu_f) \\
\mathrm{Cov}[r_t|z_{t-1}] &= \sum_{s=1}^{S} \hat{p}(s) \cdot \Big(\alpha_s \alpha_s^\top + \alpha_s \mu_f^\top B_s^\top + B_s \mu_f \alpha_s^\top + B_s(\Sigma_f + \mu_f \mu_f^\top) B_s^\top + \operatorname{diag} \sigma_s^2\Big) \\
&\quad - \left(\sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s \mu_f)\right)\left(\sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s \mu_f)\right)^\top
\end{aligned}$$

I perform an empirical experiment as follows. I set a training period of $T_{\text{train}} = 300$ days and a holding period of $H = 7$ days. At the end of the first training period and at the end of each holding period, I fit a three-regime RSFM using the past T_{train} observations. Given some $\gamma > 0$, I compute the optimal w under mean-variance utility and restrictions on short-selling using the estimated $\mathbb{E}[r_t|z_{t-1}]$ and $\mathrm{Cov}[r_t|z_{t-1}]$. I include the restrictions on short-selling to be more realistic, as it is often difficult to short on cryptocurrency trading platforms. The short-selling restrictions also imply that many of the optimal weights will be set to zero, leading to sparse portfolios, another feature of real-world investing behavior. I record the performance of the portfolio over a test period, rebalancing the portfolio every H days, and compare its performance against simpler estimates of the mean vector and covariance matrix, such as using a factor model or the empirical distribution. Practical considerations for running this experiment are in the appendix.

4.2 Results

Figure 4 shows the average returns over the test period when using the RSFM, factor model (FM) and empirical distribution (ED) to estimate the mean vector and covariance matrix of asset returns against various values of the risk aversion parameter. The figure suggests that the outperformance of the RSFM estimates to the FM and ED estimates is robust to values of the risk aversion parameter γ. The figure also suggests that over the test period, trading using RSFM estimates outperforms a momentum strategy with $H = 7$ and no short positions, i.e. buying the value-weighted portfolio of the coins in the highest quintile for one-week momentum at each rebalancing day. Figure 5 shows the results of the experiment in Section 4.1 ran with training periods of $T_{\text{train}} = 330$ (left) and $T_{\text{train}} = 365$ (right), showing the robustness of the RSFM estimates' overperformance to training period length.

Figure 4: Portfolio Performance

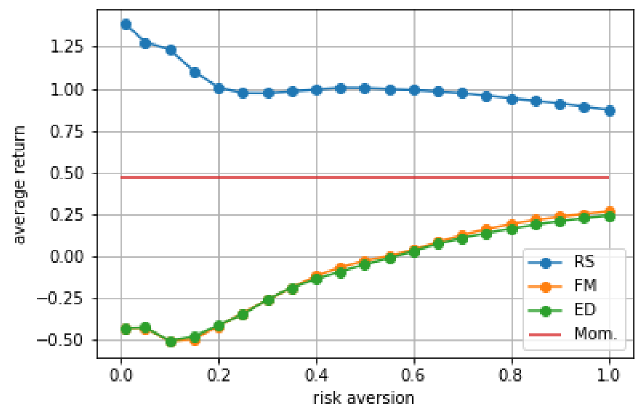
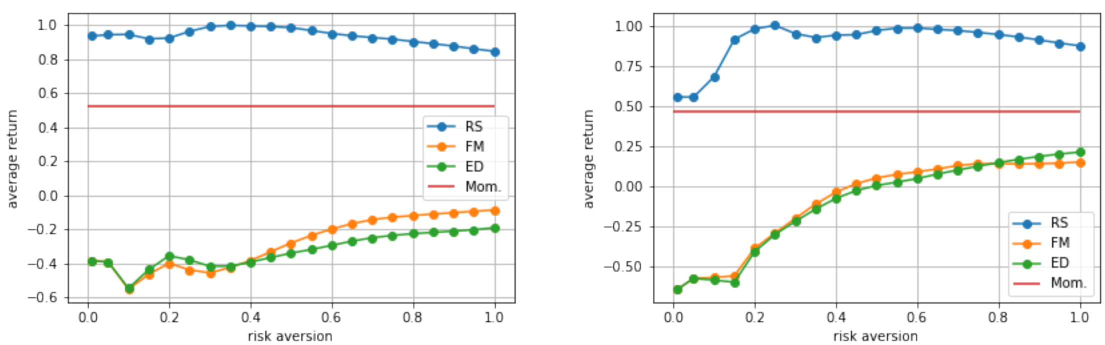


Figure 5: Additional Portfolio Performance



5 Demand Estimation

5.1 Setting

Suppose there are J heterogeneous investors. Benetton and Compiani (2020) follow the asset pricing model of Koijen and Yogo (2019) and model investor j's portfolio weights relative to an outside asset subscripted by 0 as

$$\frac{w_{t,i,j}}{w_{0,i,j}} = \exp(\alpha \cdot \text{mc}_{t,i} + \beta X_{t,i} + \gamma B_{i,j} + \lambda D_j) \cdot \varepsilon_{t,i,j}$$

where $\text{mc}_{t,i}$ is the log-market capitalization and $X_{t,i}$ is a vector of characteristics (here, a proof-of-work indicator, market beta, and four-week momentum) of coin i at time t, $B_{i,j}$ is a vector of beliefs that investor j has about coin i, D_j is a vector of demographic data for investor j, and $\varepsilon_{t,i,j}$ is a shock term capturing any unobserved factors affecting $w_{t,i,j}/w_{t,0,j}$. Benetton and Compiani (2020) use holdings data as the $w_{t,i,j}/w_{t,0,j}$ and binary variables encoding survey responses as the $B_{i,j}$. The parameter vector $(\alpha, \beta, \gamma, \lambda)$ is estimated using the generalized method of moments (GMM) with a supply-side instrument for the log-market capitalization.

To use the regime-switching framework to estimate the effect of investor beliefs on demand for cryptocurrencies, I assume the following:

- The market has S regimes, and returns have a regime-dependent factor structure. At each time t, the market regime is drawn from some true distribution p_t.
- Investors have mean-variance utilities and short-selling restrictions on their cryptocurrency portfolios, given their perception of the market regime distribution. Under this assumption, investor heterogeneity and pricing anomalies can be explained by divergences from the perceived distributions to the true distribution.
- Investors allocate a fixed amount of their wealth to an outside asset representative of non-cryptocurrency assets for the entire test period. This is so that $w_{t,i,j}/w_{t,0,j}$ can be normalized to simply $w_{t,i,j}$.

I simulate the second assumption and introduce investor heterogeneity by setting each investor's perception of the market regime distribution to be the estimated latent variable distribution exponentially tilted by the investor's individual belief bias parameters. Formally, each investor $j \in \{1, \ldots, J\}$ is given a parameter θ_j representing the investor's predisposition towards the bandwagon effect and a parameter η_j representing the investor's predisposition towards belief in high-volatility states. To simulate these representations of belief biases, I tilt the estimated latent variable distribution $\hat{p}_t$ at time t for investor j as follows:

- Draw θ_j and η_j from distributions D_θ and D_η, respectively. At each test or rebalancing day t, do the following:
  - Fit a (three-regime) RSFM using the last T_{train} observations from time $t-1$. Obtain the prediction probabilities $\hat{p}_t$ at time t and the estimated idiosyncratic risk $\hat{\sigma}_1, \ldots, \hat{\sigma}_S$.

– Let $s'_{(1)}, \ldots, s'_{(S)}$ be the regimes ordered in increasing order by their probability under $\hat{p}_t$. Let $\tilde{p}_{t,j}$ be an intermediary distribution given by

$$\tilde{p}_{t,j}(s'_{(k)}) = \frac{\hat{p}_t(s'_{(k)}) \cdot \exp(\theta_j k)}{\sum_{k'=1}^{S} \hat{p}_t(s'_{(k')}) \cdot \exp(\theta_j k')}$$

– Let $s''_{(1)}, \ldots, s''_{(S)}$ be the regimes ordered in increasing order by their average idiosyncratic risk under $(\hat{\sigma}_1, \ldots, \hat{\sigma}_S)$. Let $\hat{p}_{t,j}$ be the tilted distribution, given by

$$\hat{p}_{t,j}(s''_{(k)}) = \frac{\tilde{p}_{t,j}(s''_{(k)}) \cdot \exp(\eta_j k)}{\sum_{k'=1}^{S} \tilde{p}_{t,j}(s''_{(k')}) \cdot \exp(\eta_j k')}$$

This tilting scheme means that if θ_j is positive (negative), then the distribution will be tilted in favor of the more (less) likely regimes, assuming that the overall market sentiment about the regime distribution is proxied by $\hat{p}_t$. Analogously, if η_j is positive (negative), then the distribution will be tilted in favor of the more (less) volatile regimes, where regime volatility is proxied by the RSFM estimates of idiosyncratic risk. These belief biases are relevant and worth analyzing with regard to the data set as social media has a heavy influence on cryptocurrency price fluctuations, and these biases are present on social media. Wang et al. (2015) finds that there is a presence of the bandwagon effect on the popularity and lifecycle of Facebook posts, and Al Guindy (2021) finds that increased social media activity corresponds to greater cryptocurrency price volatility, suggesting a connection between investor attention and cryptocurrency price volatility.

Whereas Benetton and Compiani (2020) use real-world holdings data, I use the assumptions of the regime-switching framework and exponential tilting to simulate holdings data. There are several strengths to this approach. First, holdings data may be unrepresentative of all investors and prone to measurement error, and binary survey responses may not elicit or capture well the true beliefs of the investors. Second, simulated data is beneficial for understanding the demand of investors who maximize utility given their beliefs, as real-world data may include investors who, after controlling for beliefs, still trade irrationally. Third, it may be difficult to capture or quantify belief biases from surveys or real-world data.

The simulated data is created as follows. At each test or rebalancing day t, I fit a (three-regime) RSFM using the last T_{train} observations from time $t-1$. Given $\hat{p}_{t,j}$ for investor j, I compute the optimal portfolio under mean-variance utility, short-selling restrictions, and risk aversion parameter γ_j like in Section 4, but using $\hat{p}_{t,j}$ instead of $\hat{p}_t$. The values of the risk aversion parameter can be fixed or drawn from a distribution D_γ. This gives $w_{t,i,j}$ for $i \in \{1, \ldots, N\}$. By the third assumption, I normalize $w_{t,0,j}$ to 1 for all investors $j \in \{1, \ldots, J\}$, so $w_{t,i,j}$ is investor j's holdings ratio of coin i to the outside asset at time t. I set the number of investors in the market to $J = 100$ and rebalance portfolios on 16 occasions for a total of 160,000 simulated data points. This method to simulate data relies on the second and third assumptions for justification. Differing slightly from this setting, Koijen and Yogo (2019) assume mean-variance utilities across the cryptocurrency portfolio *and* the outside asset, and fix the risk aversion parameter to $\gamma = 1$.

I follow Benetton and Compiani (2020) in their choice of cryptocurrency controls: log-market capitalization, proof-of-work indicator, market beta via the CAPM, and one-week momentum of coin i at time t; I add the investors' belief bias parameters. The observation vector $X_{t,i,j}$ is then

$$X_{t,i,j} = (\mathrm{mc}_{t,i}, \mathbb{I}[\mathrm{PoW}]_i, \mathrm{MKT}_{t,i}, \mathrm{MOM}_{t,i}, \gamma_j, \theta_j, \eta_j)$$

Benetton and Compiani (2020) observe that log-market capitalization may be endogenous due to its simultaneity in determining cryptocurrency supply and demand, and instrument on the log-supply, where the supply of coin i at time t is given by $\mathrm{SUP}_{t,i} = \mathrm{MC}_{t,i}/\pi_{t,i}$. To show the relevance of the instrument, I follow the first-stage regression of Benetton and Compiani (2020), which only includes cryptocurrency characteristics:

$$\mathrm{mc}_{t,i} = \psi^\top(\mathrm{sup}_{t,i}, \mathbb{I}[\mathrm{PoW}]_i, \mathrm{MKT}_{t,i}, \mathrm{MOM}_{t,i})$$

with parameter vector ψ. The instrumented observation is

$$Z_{t,i,j} = (\mathrm{sup}_{t,i}, \mathbb{I}[\mathrm{PoW}]_i, \mathrm{MKT}_{t,i}, \mathrm{MOM}_{t,i}, \gamma_j, \theta_j, \eta_j)$$

Let $y_{t,i,j} = w_{t,i,j}/w_{t,0,j}$. The demand equation is then

$$y_{t,i,j} = \exp(\beta^\top X_{t,i,j}) \cdot \varepsilon_{t,i,j}$$

with parameter vector β and moment condition

$$\mathbb{E}[\varepsilon_{t,i,j} | Z_{t,i,j}] = 1$$

I then use the GMM to estimate β. Practical considerations for estimating β are in the appendix.

5.2 Results

Table 6 displays the results of the first-stage regression. All coefficients are significant at the 0.05 level, with most significant at the 0.01 level as well, demonstrating instrument relevance of the log-supply. Table 7 displays the GMM estimation results for various combinations of variables. In all of these settings, the coefficients on log-market capitalization and proof-of-work are negative with significance, suggesting that investors generally have a higher level of demand for coins that are small and/or do not use proof-of-work. The coefficients on the market beta and momentum are positive with significance, suggesting that investors generally have higher demand for coins with high market beta exposure and/or momentum. Since coins requiring proof-of-work tend to be larger, size and returns are negatively correlated, and momentum and returns are positively correlated, it appears that these rational simulated investors thus have higher demand for coins with higher returns. The market beta coefficient suggests that these investors have higher demand for more volatile coins, even after controlling for risk aversion.

Table 6: First-Stage Regression

Dependent Variable: mc	sup	$\mathbb{I}$[PoW]	MKT	MOM
	0.6907***	1.0091***	7.0477***	0.0601**
	(0.017)	(0.208)	(0.323)	(0.026)

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

In the results of Benetton and Compiani (2020), the coefficients on log-market capitalization and the proof-of-work indicator are positive with significance, and the coefficient on four-week momentum is positive but not significant. This may be because of everyday real-world investors typically hearing only of larger cryptocurrencies, many of which require proof-of-work; on the other hand, smaller coins tend to have higher returns, so utility-maximizing investors will demand more of those smaller coins, many of which do not require proof-of-work. Recalling that I use one-week momentum instead of four-week momentum, it appears that one-week momentum is a better predictor of demand than four-week momentum.

In all settings, the coefficient on the bandwagon effect belief bias are negative with significance. At face value, this can be interpreted as a lower base level of demand for investors who are more prone to bandwagoning. However, most of the rebalancing days used to simulate data points had the regime with the lowest average volatility as the likeliest state, so it may be the case θ is indirectly capturing beliefs about volatility, i.e. high values of θ are practically like low values of η. This can be ultimately ruled out as the coefficient on η, although not significant in most settings, has the same sign as the coefficient on θ, as it is expected to have the opposite sign under that explanation. Furthermore, the coefficient on η becomes more significant when not controlling for θ, suggesting that there is more than just the relative volatility of the likeliest state influencing demand, so the face-value interpretation has some validity for now and warrants further exploration. The coefficient on risk aversion is negative with significance in all settings, implying that more risk-averse investors have a lower base level of demand. This in turn may imply that the basket of cryptocurrencies is too risky, in their opinion, compared to the outside asset(s).

It should be emphasized that these results provide insights on cryptocurrency demand solely for utility-maximizing investors, and do not characterize cryptocurrency demand in the real world. Investors' traits and belief biases were drawn according to $D_\theta, D_\eta \sim \mathcal{U}(-1,1)$ and $D_\gamma \sim \mathcal{U}(0,1)$. Robustness checks on the underlying distribution where traits and belief biases are drawn from are in the appendix.

Table 7: Demand Estimation

	$D_\theta, D_\eta = \mathcal{U}(-1,1)$; $D_\gamma = \mathcal{U}(0,1)$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
mc	-0.244***	-0.235***	-0.243***	-0.234***	-0.227***	-0.234***	-0.226***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)
$\mathbb{I}$[PoW]	-1.436***	-1.590***	-1.429***	-1.591***	-1.600***	-1.444***	-1.592***
	(0.086)	(0.084)	(0.086)	(0.084)	(0.081)	(0.083)	(0.081)
MKT	0.836***	0.684***	0.809***	0.664***	0.630***	0.756***	0.629***
	(0.117)	(0.123)	(0.116)	(0.122)	(0.124)	(0.117)	(0.123)
MOM	0.068***	0.066***	0.068***	0.066***	0.066***	0.068***	0.066***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
γ					-0.265***	-0.234**	-0.300***
					(0.091)	(0.092)	(0.090)
θ		-0.271***		-0.267***	-0.271***		-0.265***
		(0.038)		(0.038)	(0.037)		(0.037)
η			-0.061*	-0.020		-0.064*	-0.026
			(0.036)	(0.036)		(0.036)	(0.035)

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

6 Conclusion

In this paper, I have considered four types of regime-switching factor models and derived closed-form update rules for the parameters of each of them. The assumption of Gaussian idiosyncratic risk is critical to derive the closed-form updates and view some of them as weighted linear regressions. With the market, size, and momentum as my factors, I fitted these models on recent cryptocurrency data, showing how the RSFM and α-RSFM with $S = 3$ regimes appear to be the best models for the data. I then demonstrated two applications of the regime-switching framework of the models. First, I estimated mean vectors and covariance matrices conditional on the latent-state distribution, and showed that mean-variance-optimal portfolios using these estimates outperform those using traditional estimation methods (e.g. non-regime-switching factor models and empirical distributions), as well as the common momentum strategy. Second, I simulated and estimated a model of demand for cryptocurrencies. The regime-switching and demand model framework allowed for testing the effect cryptocurrency characteristics, risk aversion, and belief biases on utility-maximizing investors' demand for cryptocurrencies, with most covariates considered having significant coefficients. Notably, the simulation suggests that higher risk aversion and bandwagoning propensity results in a lower base level of demand.

There is plentiful room for further exploration with regard to both the model formulation and application areas. As hinted in Section 3.3, future work could incorporate serial components to the models, such as regime-switching vector autoregressive models for the factor time series or regime-switching conditional heteroskedasticity models for the idiosyncratic risk part of the factor model. For portfolio selection, the regime-switching framework could be incorporated in some of the more innovative estimation methods described in Section 4.1 (e.g. robust optimization and resampling). For demand estimation, additional tilting schemes could be considered that represent other investor traits or belief biases. The tilting scheme proposed here could also motivate research into empirical methods of estimating and calibrating the tilting parameters given real-world investor data. Overall, regime-switching factor models are rich in their mathematical and statistical structure, as well as in their practical uses.

References

Al Guindy, M. (2021). Cryptocurrency price volatility and investor attention. *International Review of Economics & Finance*, 76, 556-570.

Augustyniak, M. (2014). Maximum likelihood estimation of the Markov-switching GARCH model. *Computational Statistics & Data Analysis*, 76, 61–75.

Bae, G. I., Kim, W. C., and Mulvey, J. M. (2014). Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research*, 234(2), 450-458.

Baum, L. E., Petrie, T., Soules, G. and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1), 164-171.

Benetton, M. and Compiani, G. (2020). Investors' beliefs and cryptocurrency prices. *Becker Friedman Institute Working Papers*, 2020-107.

Billio, M., Casarin R., Osuntuyi, A. (2016). Efficient Gibbs sampling for Markov switching GARCH models. *Computational Statistics & Data Analysis*, 100, 37-57.

Costa, G. and Kwon, R. H. (2019). Risk parity portfolio optimization under a markov regime-switching framework. *Quantitative Finance*, 19(3), 453-471.

Costa, G. and Kwon, R. H. (2020). A regime-switching factor model for mean-variance optimization. *Journal of Risk*, 22(4).

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Giglio, S., Maggiori, M., Stroebel, J. and Utkus, S. (2021). Five facts about beliefs and portfolios. *American Economic Review*, 111(5), 1481-1522.

Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems. M*athematics of Operations Research*, 28(1), 1–38.

Koijen, R. S. J. and Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4), 1475-1515.

Kim, C. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics*, 60, 1–22.

Kim, M.-J. and Yoo, J.-S. (1995). New index of coincident indicators: A multivariate Markov switching factor model approach. *Journal of Monetary Economics*, 36(3), 607-630.

Kuan, C.-M. (2002). *Lecture on the Markov-switching model.* Unpublished manuscript. Retrived from https://homepage.ntu.edu.tw/~ckuan/pdf/Lec-Markov_note.pdf.

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

Hamilton, J. D. and Sumsel R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1-2), 307-333.

Hubrich, S. (2017). 'Know When to Hodl 'Em, Know When to Fodl 'Em': An Investigation of Factor Based Investing in the Cryptocurrency Space. *SSRN Electronic Journal*, 10.2139/ssrn.3055498.

Liu, Y. and Tsyvinski A. (2021). Risks and returns of cryptocurrencies. *The Review of Financial Studies*, 34(6), 2689-2727.

Michaud, R. O. and Michaud, R. O. (2008). *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation.* Oxford University Press.

Nystrup, P., Madsen, H., and Lindström, E. (2018). Dynamic portfolio optimization across hidden market regimes. *Quantitative Finance*, 18(1), 83-95.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3), 425-442.

Shen, D., Urquhart, A., and Wang, P. (2019). Does twitter predict bitcoin? *Economics Letters*, 174, 118-122.

Shen, D., Urquhart A., and Wang P. (2020). A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, 34, 101248.

Urquhart, A. (2016). The inefficiency of bitcoin. *Economics Letters*, 148, 80-82.

Urquhart, A. (2018). What causes the attention of bitcoin? *Economics Letters*, 166, 40–44.

Wang, K. C., Lai, C.-M., Wang, T., and Wu, S. F. (2015). Bandwagon effect in Facebook discussion groups. *Proceedings of ASE Big Data & Social Informatics 2015*, 17.

A Appendix

A.1 Derivations for Section 2.2

A.1.1 Regime-Switching Factor Model

Let $\theta_{\text{FM}} = \big((\alpha_s)_{s=1}^S, (B_s)_{s=1}^S, (\sigma_s)_{s=1}^S\big)$. Define

$$\mathcal{L}(\theta_{\text{FM}}) = \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta]$$

Note that $(r_t | z_t = s, f) \sim \mathcal{N}(\alpha_s + B_s f_t, \operatorname{diag} \sigma_s^2)$. Let $B_{s,i}$ be the i-th row of B_s. Then

$$\begin{aligned}
\mathcal{L}(\theta_{\text{FM}}) &= \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) \left(-\frac{N\log 2\pi}{2} - \frac{\log\det\operatorname{diag}\sigma_s^2}{2} - \frac{(r_t - \alpha_s - B_s f_t)^\top (\operatorname{diag}\sigma_s^2)^{-1}(r_t - \alpha_s - B_s f_t)}{2}\right) \\
&= \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) \left(-\frac{\sum_{i=1}^N \log 2\pi}{2} - \frac{\sum_{i=1}^N \log \sigma_{s,i}^2}{2} - \sum_{i=1}^{N} \frac{(r_t - \alpha_{s,i} - B_{s,i} f_t)^2}{2\sigma_{s,i}^2}\right) \\
&= \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) \sum_{i=1}^{N} \left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_{s,i}^2}{2} - \frac{(r_{t,i} - \alpha_{s,i} - B_{s,i} f_t)^2}{2\sigma_{s,i}^2}\right) \\
&= \sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{i=1}^{N} q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_{s,i}^2}{2} - \frac{(r_{t,i} - \alpha_{s,i} - B_{s,i} f_t)^2}{2\sigma_{s,i}^2}\right) \\
&= \sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{t=1}^{T} q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_{s,i}^2}{2} - \frac{(r_{t,i} - \alpha_{s,i} - B_{s,i} f_t)^2}{2\sigma_{s,i}^2}\right)
\end{aligned}$$

Let $\beta_{s,i} = (\alpha_{s,i}, B_{s,i})^\top$, $\phi_t = (1, f_{t,1}, \ldots, f_{t,M})^\top$,, and $\tilde{\theta}_{\text{FM}} = ((\beta_s)_{s=1}^S, (\sigma_s)_{s=1}^S)$. Define

$$\tilde{\mathcal{L}}(\tilde{\theta}_{\text{FM}}) = \sum_{t=1}^{S}\sum_{i=1}^{N}\sum_{s=1}^{T} q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\sigma_{s,i}}{2} - \frac{(r_{t,i} - \beta_{s,i}^\top \phi_t)^2}{2\sigma_{s,i}^2}\right)$$

Then $\mathcal{L}(\theta_{\text{FM}}) = \tilde{\mathcal{L}}(\tilde{\theta}_{\text{FM}})$. Note that $\tilde{\mathcal{L}}(\tilde{\theta}_{\text{FM}})$ is concave, so the first-order conditions (FOCs) for $\tilde{\mathcal{L}}(\tilde{\theta}_{\text{FM}})$ are sufficient for the optimum. The FOC for $\beta_{s,i}$ is

$$\frac{\partial \tilde{\mathcal{L}}(\tilde{\theta}_{\text{FM}})}{\partial \beta_{s,i}} = \sum_{t=1}^{T} \frac{q_t(s)(r_{t,i} - \beta_{s,i}^\top \phi_t)\phi_t}{\sigma_{s,i}^2} = 0$$

$$\sum_{t=1}^{T} \left((q_t(s))^{1/2} r_{t,i} - \beta_{s,i}^\top (q_t(s))^{1/2} \phi_t\right) (q_t(s))^{1/2} \phi_t = 0$$

Define

$$\tilde{q}_s = \left((q_1(s))^{1/2}, \ldots, (q_T(s))^{1/2}\right)^\top$$
$$\tilde{r}_{s,i} = \left((q_1(s))^{1/2} r_{1,i}, \ldots, (q_T(s))^{1/2} r_{T,i}\right)^\top$$
$$\Phi_s = \begin{bmatrix} (q_1(s))^{1/2} & (q_1(s))^{1/2} f_{1,1} & \cdots & (q_1(s))^{1/2} f_{1,M} \\ \vdots & \vdots & \ddots & \vdots \\ (q_T(s))^{1/2} & (q_T(s))^{1/2} f_{T,1} & \cdots & (q_T(s))^{1/2} f_{T,M} \end{bmatrix}$$

In the ordinary least squares setting of $y_t = \beta^\top x_t + \varepsilon_t$, the FOC for β is $\sum_{t=1}^T (y_t - \beta^\top x_t) x_t = 0$. Matching terms with the RSFM setting gives the M-step update for $\beta_{s,i}$ as

$$\beta^\star_{s,i} = (\Phi_s^\top \Phi_s)^{-1} \Phi_s^\top \tilde{r}_{s,i}$$

The FOC for $\sigma_{s,i}$ is

$$\frac{\partial \tilde{\mathcal{L}}(\tilde{\theta}_{\mathrm{FM}})}{\partial \sigma_{s,i}} = \sum_{t=1}^T \frac{1}{\sigma_{s,i}} \left(\frac{q_t(s)(r_{t,i} - \beta_{s,i}^\top \phi_t)^2}{2\sigma_{s,i}^2} - \frac{q_t(s)}{2} \right) = 0$$
$$\frac{\sum_{t=1}^T q_t(s)(r_{t,i} - \beta_{s,i}^\top \phi_t)^2}{\sigma_{s,i}^2} = \sum_{t=1}^T q_t(s)$$

Solving the FOC for $\sigma_{s,i}$ gives the M-step update for $\sigma_{s,i}$ as

$$\sigma^\star_{s,i} = \frac{\|\tilde{r}_{s,i} - \Phi_s \beta^\star_{s,i}\|_2}{\|\tilde{q}_s\|_2}$$

A.1.2 Alpha-Restricted Regime-Switching Factor Model

Let $\theta_{\mathrm{FM}} = \left(\alpha, (B_s)_{s=1}^S, (\sigma_s)_{s=1}^S\right)$. Define

$$\mathcal{L}_\alpha(\theta_{\mathrm{FM}}) = \sum_{t=1}^T \sum_{s=1}^S q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta]$$

Note that $(r_t | z_t = s, f) \sim \mathcal{N}(\alpha + B_s f_t, \operatorname{diag} \sigma_s^2)$. Let $B_{s,i}$ be the i-th row of B_s. Then

$$\mathcal{L}_\alpha(\theta_{\mathrm{FM}}) = \sum_{i=1}^N \sum_{s=1}^S \sum_{t=1}^T q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_{s,i}^2}{2} - \frac{(r_{t,i} - \alpha_i - B_{s,i} f_t)^2}{2\sigma_{s,i}^2} \right)$$

Note that $\mathcal{L}_\alpha(\theta_{\text{FM}})$ is concave in each parameter. The FOC for α_i is

$$\frac{\partial \mathcal{L}_\alpha(\theta_{\text{FM}})}{\partial \alpha_i} = \sum_{s=1}^{S}\sum_{t=1}^{T} \frac{q_t(s)\cdot(r_{t,i}-\alpha_i - B_{s,i}f_t)}{\sigma_{s,i}^2} = 0$$

which can be writen as

$$\begin{aligned}
\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(r_{t,i}-\alpha_i - B_{s,i}f_t) = 0 \\
\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)\cdot\alpha_i &= \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(r_{t,i} - B_{s,i}f_t) \\
\alpha_i \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s) &= \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)\cdot r_{t,i} - \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(B_{s,i}f_t) \\
\alpha_i \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) &= \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s)\cdot r_{t,i} - \sum_{s=1}^{S}\sum_{t=1}^{T} B_{s,i}(q_t(s)\cdot f_t) \\
T\alpha_i &= \sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} B_{s,i}\left(\sum_{t=1}^{T} q_t(s)\cdot f_t\right)
\end{aligned}$$

The FOC for $B_{s,i}$ is

$$\frac{\partial \mathcal{L}_\alpha(\theta_{\text{FM}})}{\partial B_{s,i}} = \sum_{t=1}^{T} \frac{q_t(s)\cdot(r_{t,i}-\alpha_i - B_{s,i}f_t)f_t}{\sigma_{s,i}^2} = 0$$

which can be written as

$$\begin{aligned}
\sum_{t=1}^{T} q_t(s)\cdot(r_{t,i}-\alpha_i - B_{s,i}f_t)f_t = 0 \\
\sum_{t=1}^{T} q_t(s)\cdot\alpha_i\cdot f_t &= \sum_{t=1}^{T} q_t(s)\cdot(r_{t,i} - B_{s,i}f_t)f_t \\
\alpha_i \sum_{t=1}^{T} q_t(s)\cdot f_t &= \sum_{t=1}^{T} q_t(s)\cdot r_{t,i}\cdot f_t - \sum_{t=1}^{T} q_t(s)\cdot(B_{s,i}f_t)f_t
\end{aligned}$$

Define

$$\begin{aligned}
q_s &= (q_1(s),\ldots,q_T(s))^\top \\
r_{s,i} &= (q_1(s)\cdot r_{1,i},\ldots,q_T(s)\cdot r_{T,1})^\top
\end{aligned}$$

$$F = \begin{bmatrix} f_{1,1} & \cdots & f_{1,M} \\ \vdots & \ddots & \vdots \\ f_{T,1} & \cdots & f_{T,M} \end{bmatrix}$$

$$\tilde{F}_s = \begin{bmatrix} (q_1(s))^{1/2} f_{1,1} & \cdots & (q_1(s))^{1/2} f_{1,M} \\ \vdots & \ddots & \vdots \\ (q_1(s))^{1/2} f_{T,1} & \cdots & (q_T(s))^{1/2} f_{T,M} \end{bmatrix}$$

Then the FOC for α_i can be written as

$$T\alpha_i = \sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} B_{s,i}(F^\top q_s)$$

and the FOC for $B_{s,i}$ can be written as

$$\alpha_i(F^\top q_s) = F^\top r_{s,i} - \tilde{F}_s^\top \tilde{F}_s B_{s,i}$$

Solving for $B_{s,i}$ gives

$$B_{s,i} = (\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha_i q_s)$$

Substituting this expression into the FOC for $\alpha_{s,i}$ gives

$$\begin{aligned} T\alpha_i &= \sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} ((\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha_i q_s))^\top (F^\top q_s) \\ &= \sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} (F^\top (r_{s,i} - \alpha_i q_s))^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s) \\ &= \sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} (F^\top r_{s,i})^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s) + \alpha_i \sum_{s=1}^{S} (F^\top q_s)^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s) \end{aligned}$$

Solving for α_i gives the M-step updates for α_i and $B_{s,i}$ as

$$\begin{aligned} \alpha_i^\star &= \frac{\sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} (F^\top r_{s,i})^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)}{T - \sum_{s=1}^{S} (F^\top q_s)^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1} (F^\top q_s)} \\ B_{s,i}^\star &= (\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha_i^\star q_s) \end{aligned}$$

The FOC for $\sigma_{s,i}$ is

$$\frac{\partial \mathcal{L}_\alpha(\theta_{\mathrm{FM}})}{\partial \sigma_{s,i}} = \sum_{t=1}^{T} \frac{1}{\sigma_{s,i}} \left(\frac{q_t(s)(r_{t,i} - \alpha_i - B_{s,i} f_t)^2}{2\sigma_{s,i}^2} - \frac{q_t(s)}{2} \right) = 0$$

$$\frac{\sum_{t=1}^{T} q_t(s)(r_{t,i} - \alpha_i - B_{s,i} f_t)^2}{\sigma_{s,i}^2} = \sum_{t=1}^{T} q_t(s)$$

Solving the FOC for $\sigma_{s,i}$ gives the M-step update for $\sigma_{s,i}$ as

$$\sigma_{s,i}^\star = \left(\frac{\sum_{t=1}^{T} q_t(s)(r_{t,i} - \alpha_i^\star - B_{s,i}^{\star\top} f_t)^2}{\sum_{t=1}^{T} q_t(s)} \right)^{1/2}$$

A.1.3 Sigma-Restricted Regime-Switching Factor Model

Let $\theta_{\mathrm{FM}} = \big((\alpha_s)_{s=1}^{S}, (B_s)_{s=1}^{S}, \sigma_s \big)$. Define

$$\mathcal{L}_\sigma(\theta_{\mathrm{FM}}) = \sum_{t=1}^{T} \sum_{s=1}^{S} q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta]$$

Note that $(r_t | z_t = s, f) \sim \mathcal{N}(\alpha_s + B_s f_t, \operatorname{diag} \sigma^2)$. Let $B_{s,i}$ be the i-th row of B_s. Then

$$\mathcal{L}_\sigma(\theta_{\mathrm{FM}}) = \sum_{i=1}^{N} \sum_{s=1}^{S} \sum_{t=1}^{T} q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_i^2}{2} - \frac{(r_{t,i} - \alpha_{s,i} - B_{s,i} f_t)^2}{2\sigma_i^2} \right)$$

Let $\beta_{s,i} = (\alpha_{s,i}, B_{s,i})^\top$, $\phi_t = (1, f_{t,1}, \ldots, f_{t,M})^\top$,, and $\tilde{\theta}_{\mathrm{FM}} = ((\beta_s)_{s=1}^{S}, (\sigma_s)_{s=1}^{S})$. Define

$$\tilde{\mathcal{L}}_\sigma(\tilde{\theta}_{\mathrm{FM}}) = \sum_{t=1}^{S} \sum_{i=1}^{N} \sum_{s=1}^{T} q_t(s) \left(-\frac{\log 2\pi}{2} - \frac{\sigma_{s,i}}{2} - \frac{(r_{t,i} - \beta_{s,i}^\top \phi_t)^2}{2\sigma_{s,i}^2} \right)$$

Then $\mathcal{L}_\sigma(\theta_{\mathrm{FM}}) = \tilde{\mathcal{L}}_\sigma(\tilde{\theta}_{\mathrm{FM}})$. The FOC for $\beta_{s,i}$ is

$$\frac{\partial \tilde{\mathcal{L}}_\sigma(\tilde{\theta}_{\mathrm{FM}})}{\partial \beta_{s,i}} = \sum_{t=1}^{T} \frac{q_t(s)(r_{t,i} - \beta_{s,i}^\top \phi_t)\phi_t}{\sigma_{s,i}^2} = 0$$

$$\sum_{t=1}^{T} \left((q_t(s))^{1/2} r_{t,i} - \beta_{s,i}^\top (q_t(s))^{1/2} \phi_t \right) (q_t(s))^{1/2} \phi_t = 0$$

so by Section A.1.1, the M-step update for $\beta_{s,i}$ is

$$\beta_{s,i}^\star = (\Phi_s^\top \Phi_s)^{-1} \Phi_s^\top \tilde{r}_{s,i}$$

The FOC for σ_i is

$$\frac{\partial \tilde{\mathcal{L}}_\sigma(\tilde{\theta}_{\text{FM}})}{\partial \sigma_i} = \sum_{s=1}^{S}\sum_{t=1}^{T} \frac{1}{\sigma_i}\left(\frac{q_t(s)(r_{t,i}-\beta_{s,i}^\top \phi_t)^2}{2\sigma_i^2} - \frac{q_t(s)}{2}\right) = 0$$

$$\frac{\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(r_{t,i}-\beta_{s,i}^\top \phi_t)^2}{\sigma_i^2} = \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s) = T$$

Solving the FOC for $\sigma_{s,i}$ gives the M-step update for σ_i as

$$\sigma_i^\star = \left(\frac{\sum_{s=1}^{S} \|\tilde{r}_{s,i} - \Phi_s \beta_{s,i}^\star\|_2^2}{T}\right)^{1/2}$$

A.1.4 Alpha- and Sigma-Restricted Regime-Switching Factor Model

Let $\theta_{\text{FM}} = \big(\alpha, (B_s)_{s=1}^{S}, \sigma_s\big)$. Define

$$\mathcal{L}_{\alpha\sigma}(\theta_{\text{FM}}) = \sum_{t=1}^{T}\sum_{s=1}^{S} q_t(s) \log \mathbb{P}[r_t | z_t = s, f, \theta]$$

Note that $(r_t | z_t = s, f) \sim \mathcal{N}(\alpha + B_s f_t, \operatorname{diag} \sigma^2)$. Let $B_{s,i}$ be the i-th row of B_s. Then

$$\mathcal{L}_{\alpha\sigma}(\theta_{\text{FM}}) = \sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)\left(-\frac{\log 2\pi}{2} - \frac{\log \sigma_i^2}{2} - \frac{(r_{t,i} - \alpha_i - B_{s,i} f_t)^2}{2\sigma_i^2}\right)$$

The FOCs for α_i and $B_{s,i}$ are

$$\frac{\partial \mathcal{L}_\alpha(\theta_{\text{FM}})}{\partial \alpha_i} = \sum_{s=1}^{S}\sum_{t=1}^{T} \frac{q_t(s)\cdot(r_{t,i}-\alpha_i - B_{s,i} f_t)}{\sigma_{s,i}^2} = 0$$

$$\frac{\partial \mathcal{L}_\alpha(\theta_{\text{FM}})}{\partial B_{s,i}} = \sum_{s=1}^{S}\sum_{t=1}^{T} \frac{q_t(s)\cdot(r_{t,i}-\alpha_i - B_{s,i} f_t) f_t}{\sigma_{s,i}^2} = 0$$

so by Section A.1.2, the M-step updates for α_i and $B_{s,i}$ are

$$\alpha_i^\star = \frac{\sum_{t=1}^{T} r_{t,i} - \sum_{s=1}^{S} (F^\top r_{s,i})^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1}(F^\top q_s)}{T - \sum_{s=1}^{S}(F^\top q_s)^\top (\tilde{F}_s^\top \tilde{F}_s)^{-1}(F^\top q_s)}$$

$$B_{s,i}^\star = (\tilde{F}_s^\top \tilde{F}_s)^{-1} F^\top (r_{s,i} - \alpha_i^\star q_s)$$

The FOC for σ_i is

$$\frac{\partial \tilde{\mathcal{L}}_\sigma(\tilde{\theta}_{\mathrm{FM}})}{\partial \sigma_i} = \sum_{s=1}^{S}\sum_{t=1}^{T} \frac{1}{\sigma_i}\left(\frac{q_t(s)(r_{t,i}-\alpha_i - B_{s,i}f_t)^2}{2\sigma_i^2} - \frac{q_t(s)}{2}\right) = 0$$

$$\frac{\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(r_{t,i}-\alpha_i - B_{s,i}f_t)^2}{\sigma_i^2} = \sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s) = T$$

Solving the FOC for σ_i gives the M-step for σ_i as

$$\sigma_i^\star = \left(\frac{\sum_{s=1}^{S}\sum_{t=1}^{T} q_t(s)(r_{t,i}-\alpha_i^\star - B_{s,i}^{\star\top} f_t)^2}{T}\right)^{1/2}$$

A.2 Derivations for Section 4.1

Let $\mu_f = \mathbb{E}[f_t]$, $\Sigma_f = \mathrm{Cov}[f_t]$, and z_t have transition matrix P. I start with the RSFM equation

$$r_t = \sum_{s=1}^{S} (\alpha_s + B_s f_t + \varepsilon_{s,t}) \cdot \mathbb{I}[z_t = s]$$

Since $\varepsilon_{s,t} \sim \mathcal{N}(0, \operatorname{diag}\sigma_s^2)$ and $\varepsilon_{s,t} \perp f_t$ by assumption,

$$\begin{aligned}
\mathbb{E}[r_t|z_{t-1}=s] &= \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|z_{t-1}=s]\cdot \mathbb{E}[r_t|z_t = s'] \\
&= \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|z_{t-1}=s]\cdot \mathbb{E}\left[\sum_{\varsigma=1}^{S}(\alpha_\varsigma + B_\varsigma f_t + \varepsilon_{\varsigma,t})\cdot \mathbb{I}[z_t=\varsigma]\middle| z_t = s'\right] \\
&= \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|z_{t-1}=s]\cdot \mathbb{E}[\alpha_{s'} + B_{s'} f_t + \varepsilon_{s',t}] = \sum_{s'=1}^{S} P(s,s')\cdot(\alpha_{s'} + B_{s'}\mu_f)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}[r_t r_t^\top|z_{t-1}=s] \\
&\quad= \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|z_{t-1}=s]\cdot \mathbb{E}[r_t r_t^\top|z_t = s'] \\
&\quad= \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|z_{t-1}=s] \\
&\qquad \cdot \mathbb{E}\left[\left(\sum_{\varsigma=1}^{S}(\alpha_\varsigma + B_\varsigma f_t + \varepsilon_{\varsigma,t})\cdot \mathbb{I}[z_t=\varsigma]\right)\left(\sum_{\varsigma=1}^{S}(\alpha_\varsigma + B_\varsigma f_t + \varepsilon_{\varsigma,t})\cdot \mathbb{I}[z_t=\varsigma]\right)^\top \middle| z_t = s'\right]
\end{aligned}$$

$$= \sum_{s'=1}^{S} \mathbb{P}[z_t = s' | z_{t-1} = s] \cdot \mathbb{E}\Big[(\alpha_{s'} + B_{s'} f_t + \varepsilon_{s',t})(\alpha_{s'} + B_{s'} f_t + \varepsilon_{s',t})^\top\Big]$$

$$= \sum_{s'=1}^{S} P(s,s') \cdot \Big(\alpha_{s'}\alpha_{s'}^\top + \alpha_{s'}\mu_f^\top B_{s'}^\top + B_{s'}\mu_f\alpha_{s'}^\top + B_{s'}(\mathbb{E}[f_t f_t^\top])B_{s'}^\top + \mathbb{E}[\varepsilon_{s',t}\varepsilon_{s',t}^\top]\Big)$$

$$= \sum_{s'=1}^{S} P(s,s') \cdot \Big(\alpha_{s'}\alpha_{s'}^\top + \alpha_{s'}\mu_f^\top B_{s'}^\top + B_{s'}\mu_f\alpha_{s'}^\top + B_{s'}(\Sigma_f + \mu_f\mu_f^\top)B_{s'}^\top + \operatorname{diag}\sigma_s^2\Big)$$

Then

$$\begin{aligned}
&\operatorname{Cov}[r_t | z_{t-1} = s] \\
&\quad = \mathbb{E}[r_t r_t^\top | z_{t-1} = s] - \mathbb{E}[r_t | z_{t-1} = s]\mathbb{E}[r_t | z_{t-1} = s]^\top \\
&\quad = \sum_{s'=1}^{S} P(s,s') \cdot \Big(\alpha_{s'}\alpha_{s'}^\top + \alpha_{s'}\mu_f^\top B_{s'}^\top + B_{s'}\mu_f\alpha_{s'}^\top + B_{s'}(\Sigma_f + \mu_f\mu_f^\top)B_{s'}^\top + \operatorname{diag}\sigma_s^2\Big) \\
&\qquad - \left(\sum_{s'=1}^{S} P(s,s') \cdot (\alpha_{s'} + B_{s'}\mu_f)\right)\left(\sum_{s'=1}^{S} P(s,s') \cdot (\alpha_{s'} + B_{s'}\mu_f)\right)^\top
\end{aligned}$$

Let $z_{t-1} \sim q$ and $\hat{p}^\top = q^\top P$. Then

$$\begin{aligned}
\mathbb{E}[r_t | z_{t-1}] &= \sum_{s=1}^{S} \mathbb{P}[z_{t-1} = s] \cdot \mathbb{E}[r_t | z_{t-1} = s] \\
&= \sum_{s=1}^{S} q(s) \sum_{s'=1}^{S} P(s,s') \cdot (\alpha_{s'} + B_{s'}\mu_f) = \sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s\mu_f)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}[r_t r_t^\top | z_{t-1}] \\
&\quad = \sum_{s=1}^{S} \mathbb{P}[z_{t-1} = s] \cdot \mathbb{E}[r_t r_t^\top | z_{t-1} = s] \\
&\quad = \sum_{s=1}^{S} q(s) \sum_{s'=1}^{S} P(s,s') \cdot \Big(\alpha_{s'}\alpha_{s'}^\top + \alpha_{s'}\mu_f^\top B_{s'}^\top + B_{s'}\mu_f\alpha_{s'}^\top + B_{s'}(\Sigma_f + \mu_f\mu_f^\top)B_{s'}^\top + \operatorname{diag}\sigma_s^2\Big) \\
&\quad = \sum_{s=1}^{S} \hat{p}(s) \cdot \Big(\alpha_s\alpha_s^\top + \alpha_s\mu_f^\top B_s^\top + B_s\mu_f\alpha_s^\top + B_s(\Sigma_f + \mu_f\mu_f^\top)B_s^\top + \operatorname{diag}\sigma_s^2\Big)
\end{aligned}$$

Then

$$
\begin{aligned}
\operatorname{Cov}[r_t|z_{t-1}] &= \mathbb{E}[r_t r_t^\top|z_{t-1}] - \mathbb{E}[r_t|z_{t-1}]\mathbb{E}[r_t|z_{t-1}]^\top \\
&= \sum_{s=1}^{S} \hat{p}(s) \cdot \Big(\alpha_s\alpha_s^\top + \alpha_s\mu_f^\top B_s^\top + B_s\mu_f\alpha_s^\top + B_s(\Sigma_f + \mu_f\mu_f^\top)B_s^\top + \operatorname{diag}\sigma_s^2\Big) \\
&\quad - \left(\sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s\mu_f)\right)\left(\sum_{s=1}^{S} \hat{p}(s) \cdot (\alpha_s + B_s\mu_f)\right)^\top
\end{aligned}
$$

A.3 Practical Considerations

A.3.1 Working with Log-Densities and Log-Probabilities

One problem that arises in practice when computing the forward and backward densities, as well as the inference probabilities, is that these quantities may be very small. To resolve this issue, compute the log-densities or log-probabilities. When computing the log-densities or log-probabilities, it may be necessary to compute an expression of the form

$$
\log \sum_{s=1}^{S} x_s
$$

where the x_s may be very small, thus making the expression intractable at first glance. To compute this expression, I use the log-sum-exp function, defined as

$$
\begin{aligned}
\operatorname{LSE}(x_1, \ldots, x_S) &= \log \sum_{s=1}^{S} \exp x_s \\
&= x_{\max} + \log \sum_{s=1}^{S} \exp(x_s - x_{\max}) = x_{\max} + \operatorname{LSE}(x_1 - x_{\max}, \ldots, x_S - x_{\max})
\end{aligned}
$$

where $x_{\max} = \max\{x_1, \ldots, x_S\}$. Note that

$$
\begin{aligned}
\log \sum_{s=1}^{S} x_s &= \operatorname{LSE}(\log x_1, \ldots, \log x_S) \\
&= \log x_{\max} + \log \sum_{s=1}^{S} \frac{x_s}{x_{\max}} = \log x_{\max} + \operatorname{LSE}(\log x_1 - \log x_{\max}, \ldots, \log x_S - \log x_{\max})
\end{aligned}
$$

which is tractable in most cases. For example, the filtering probability for state s at time t is

$$\mathbb{P}[z_t = s|\mathcal{R}_t, \theta] = \frac{\mathbb{P}[z_t = s|\mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t|z_t = s, \mathcal{R}_{t-1}, \theta]}{\sum_{s'=1}^{S} \mathbb{P}[z_t = s'|\mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t|z_t = s', \mathcal{R}_{t-1}, \theta]}$$

Let $x_s = \mathbb{P}[z_t = s'|\mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t|z_t = s', \mathcal{R}_{t-1}, \theta]$. Then

$$\log x_s = \log \mathbb{P}[z_t = s'|\mathcal{R}_{t-1}, \theta] + \log \mathbb{P}[r_t|z_t = s', \mathcal{R}_{t-1}, \theta]$$

is tractable, so

$$\begin{aligned}
\log \mathbb{P}[z_t = s|\mathcal{R}_t, \theta] &= \log \mathbb{P}[z_t = s|\mathcal{R}_{t-1}, \theta + \log \mathbb{P}[r_t|z_t = s, \mathcal{R}_{t-1}, \theta] \\
&\qquad - \log \sum_{s'=1}^{S} \mathbb{P}[z_t = s'|\mathcal{R}_{t-1}, \theta] \cdot \mathbb{P}[r_t|z_t = s', \mathcal{R}_{t-1}, \theta] \\
&= \log \mathbb{P}[z_t = s|\mathcal{R}_{t-1}, \theta + \log \mathbb{P}[r_t|z_t = s, \mathcal{R}_{t-1}, \theta] \\
&\qquad - \log x_{\max} - \mathrm{LSE}(\log x_1 - \log x_{\max}, \ldots, \log x_S - \log x_{\max})
\end{aligned}$$

A.3.2 Portfolio Selection

The optimal portfolio is given by

$$\begin{aligned}
w^\star &= \arg \max_{w \in \mathbb{R}^N} \left(\mu^\top w - \frac{\gamma}{2} w^\top \Sigma w \right) \text{ s.t. } 1_N^\top w = 1,\ w \geq 0_N \\
&= \arg \min_{w \in \mathbb{R}^N} \left(\frac{\gamma}{2} w^\top \Sigma w - \mu^\top w \right) \text{ s.t. } 1_N^\top w = 1,\ w \geq 0_N
\end{aligned}$$

This is a convex optimization problem over a simplex, so exponential gradient descent can efficiently compute an approximately-optimal portfolio: choose a step size $\eta \geq 0$, set $w_0 = 1_N/N$, and update until a termination condition is reached according to

$$\begin{aligned}
g_k &= \gamma \Sigma w_k - \mu \\
Z_k &= \sum_{i=1}^{n} w_{k,i} \cdot \exp(-\eta g_{k,i}) \\
w_{k+1,i} &= \frac{w_{k,i} \cdot \exp(-\eta g_{k,i})}{Z_k}
\end{aligned}$$

The investor holds this portfolio until the next test or rebalancing day, and I track its returns.

A.3.3 Demand Estimation

I start with the demand equation

$$\frac{w_{t,i,j}}{w_{t,0,j}} = \exp\Big(\beta^\top X_{t,i,j}\Big) \cdot \varepsilon_{t,i,j}$$

and moment condition

$$\mathbb{E}[\varepsilon_{t,i,j}|Z_{t,i,j}] = 1$$

By the demand equation, the moment condition becomes

$$\mathbb{E}\left[\left(\frac{w_{t,i,j}/w_{t,0,j}}{\exp(\beta^\top X_{t,i,j})} - 1\right)\Bigg| Z_{t,i,j}\right] = 0$$

Let $y_{t,i,j} = w_{t,i,j}/w_{t,0,j}$. Define

$$g(\beta, y_{t,i,j}, X_{t,i,j}, Z_{t,i,j}) = \left(\frac{y_{t,i,j}}{\exp(\beta^\top X_{t,i,j})} - 1\right) Z_{t,i,j}$$

Statistical efficiency issues are negligible as the data set is large, so with I as the weighting matrix, a GMM estimator for $\hat{\beta}$ is

$$\hat{\beta} = \arg\min_{\beta} \frac{1}{2}\left\| \frac{1}{TNJ} \sum_{t,i,j} g(\beta, y_{t,i,j}, X_{t,i,j}, Z_{t,i,j}) \right\|_2^2$$

The GMM estimator $\hat{\beta}$ is consistent. Define

$$\hat{G} = \hat{\mathbb{E}}[\nabla g(\beta, y_{t,i,j}, X_{t,i,j}, Z_{t,i,j})] = \frac{1}{TNJ} \sum_{t,i,j} \left(-\frac{y_{t,i,j}}{\exp(\beta^\top X_{t,i,j})}\right) Z_{t,i,j} X_{t,i,j}^\top$$

$$\hat{\Omega} = \hat{\mathbb{E}}[g(\beta, y_{t,i,j}, X_{t,i,j}, Z_{t,i,j}) g(\beta, y_{t,i,j}, X_{t,i,j}, Z_{t,i,j})^\top] = \frac{1}{TNJ} \sum_{t,i,j} \left(\frac{y_{t,i,j}}{\exp(\beta^\top X_{t,i,j})} - 1\right)^2 Z_{t,i,j} Z_{t,i,j}^\top$$

Then the estimated covariance matrix of $\hat{\beta}$ is

$$\widehat{\mathrm{Cov}}[\hat{\beta}] = \frac{1}{TNJ} (\hat{G}^\top \hat{G})^{-1} (\hat{G}^\top \hat{\Omega} \hat{G}) (G^\top \hat{G})^{-1}$$

which can then be used to compute t-statistics for the coefficients of β.

A.4 Additional Tables

A.4.1 Model Selection

Table 8: Bayesian Information Criteria

	Number of Regimes				
Model	1	2	3	4	5
RSFM	270.31	262.97	261.95	263.57	263.77
α-RSFM	270.31	262.69	261.58	262.39	263.85
σ-RSFM	270.31	269.64	270.43	271.73	273.06
$\alpha\sigma$-RSFM	270.31	269.07	269.27	269.68	270.46

Note: BIC values are in thousands.

A.4.2 Demand Estimation

Table 9: Demand Estimation II

	$D_\theta, D_\eta = \mathcal{N}(0, 0.5)$ $D_\gamma = \text{Exp}(0.25)$	$D_\theta, D_\eta = \mathcal{N}(0, 0.5)$ $D_\gamma = \mathcal{U}(0, 1)$	$D_\theta, D_\eta = \mathcal{U}(-1, 1)$ $D_\gamma = \text{Exp}(0.25)$	$D_\theta, D_\eta = \mathcal{U}(-1, 1)$ $D_\gamma = \mathcal{U}(0, 1)$
mc	-0.249***	-0.223***	-0.260***	-0.226***
	(0.008)	(0.007)	(0.008)	(0.007)
$\mathbb{I}$[PoW]	-1.488***	-1.815***	-1.286***	-1.592***
	(0.107)	(0.088)	(0.010)	(0.081)
MKT	1.006***	0.594***	1.209***	0.629***
	(0.152)	(0.125)	(0.151)	(0.123)
MOM	0.067***	0.069***	0.062***	0.066***
	(0.005)	(0.004)	(0.005)	(0.004)
γ	-0.293***	-0.339***	-0.296***	-0.300***
	(0.092)	(0.084)	(0.106)	(0.090)
θ	-0.216***	-0.301***	-0.150***	-0.265***
	(0.058)	(0.047)	(0.046)	(0.037)
η	-0.007	-0.034	-0.016	-0.026
	(0.053)	(0.045)	(0.042)	(0.035)

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.