# Essays on Portfolio Selection, Continuous-time Analysis, and Market Incompleteness

by

Yixuan Li

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Economics

Waterloo, Ontario, Canada, 2023

## Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Xin Jin
Associate Professor, School of Economics
Shanghai University of Finance and Economics

Supervisor(s): Tao Chen
Associate Professor, Department of Economics
University of Waterloo

Internal Member: Francisco Gonzalez
Professor, Department of Economics
University of Waterloo

Thomas Parker
Associate Professor, Department of Economics
University of Waterloo

Internal-External Member: Ruodu Wang
Professor, Department of Statistics and Actuarial Science
University of Waterloo

## Author's Declaration

This thesis consists of material all of which I co-authored: see Statement of Contributions included in the thesis.

This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Statement of Contributions

The first chapter is coauthored with Dr. Jiawen Xu from University of Shanghai for Science and Technology, Dr. Kai Liu from University of Prince Edward Island, and my supervisor Dr. Tao Chen from University of Waterloo. The second chapter is coauthored with my supervisor Dr. Tao Chen and Dr. Renfang Tian from King's University College at Western University. The third chapter is coauthored with my supervisor Dr. Tao Chen. In all chapters, I have made major contribution to the development of research questions, the collection (or generation) and analysis of data, as well as the drafting and submission of manuscripts.

## Abstract

This thesis consists of three self-contained essays evaluating topics in portfolio selection, continuous-time analysis, and market incompleteness.

The two opposing investment strategies, diversification and concentration, have often been directly compared. Despite the less debate regarding Markowitz's approach as the benchmark for diversification, the precise meaning of concentration in portfolio selection remains unclear. Chapter 1, coauthored with Jiawen Xu, Kai Liu, and Tao Chen, offers a novel definition of concentration, along with an extreme value theory-based estimator for its implementation. When overlaying the performances derived from diversification (in Markowitz's sense) and concentration (in our definition), we find an implied risk threshold, at which the two polar investment strategies reconcile – diversification has a higher expected return in situations where risk is below the threshold, while concentration becomes the preferred strategy when the risk exceeds the threshold. Different from the conventional concave shape, the estimated frontier resembles the shape of a seagull, which is piecewise concave. Further, taking the equity premium puzzle as an example, we demonstrate how the family of frontiers nested inbetween the estimated curves can provide new perspectives for research involving market portfolios.

Parametric continuous-time analysis for stochastic processes often entails the generalization of a predefined discrete formulation to a continuous-time limit. However, unknown convergence rates of the frequency-dependent parameters can destabilize the continuous-time generalization and cause modelling discrepancy, which in turn leads to unreliable estimation and forecast. To circumvent this discrepancy, Chapter 2, coauthored with Tao Chen and Renfang Tian, proposes a simple solution based on functional data analysis and truncated Taylor series expansions. It is demonstrated through a simulation that our proposed method is superior in both fitting and forecasting continuous-time stochastic processes compared with parametric methods that encounter troubles uncovering the true underlying processes.

When the markets are incomplete, perfect risk sharing is impossible and the law of one price no longer guarantees the uniqueness of the stochastic discount factor (SDF), resulting in a set of admissible SDFs, which complicates the study of financial market equilibrium, portfolio optimization, and derivative securities. Chapter 3, coauthored with Tao Chen, proposes a discrete-time econometric framework for estimating this set of SDFs, where the market is incomplete in that there are extra states relative to the existing assets. We show that the estimated incomplete market SDF set has a unique boundary point, and shrinks to this point only when the market completes. This property allows us to develop a novel measure for market incompleteness based upon the Wasserstein metric, which estimates the least distance between the probability distributions of the complete and incomplete market SDFs. To facilitate the parametrization of market incompleteness for implementation, we then consider in detail a continuous-time framework, in which the incompleteness specifically arises from stochastic jumps in asset prices, and we demonstrate that the theoretical results developed under the discrete-time setting still hold true. Furthermore, we study the evolution of market incompleteness in four of the world's major stock markets, namely those in China, Japan, the United Kingdom, and the United States. Our findings indicate that an increase in market incompleteness is usually caused by financial crises or policy changes that raise the likelihood of unanticipated risks.

## Acknowledgements

I would like to express my sincere appreciation to my thesis committee comprising Professors Tao Chen, Francisco Gonzalez, Thomas Parker, and Ruodu Wang for their excellent guidance, valuable knowledge, and endless support. The completion of this thesis would not have been possible without the countless insightful discussions with them. I owe my deep gratitude to Professor Tao Chen, my supervisor, for introducing me to the potential of econometrics in solving real-life issues, where I have found my passion and gained profound life-changing experiences. His unreserved training, coaching and encouragement have motivated me to never falter in the face of adversity, and continue exploring the world in a creative and passionate manner.

Without the knowledge and expertise of all faculty, staff, and my colleagues, I could not have been able to embark on this journey. My deepest appreciation goes out to all members of the Economics department for their selfless help and encouragement throughout my PhD. They have made my experience at the University of Waterloo memorable and meaningful.

In closing, I would like to thank my family for their love and support. I cannot express how grateful I am to my parents, Wu Li and Wenli Zhu, who have always believed in me and supported me in all my endeavours. Furthermore, I am extremely grateful to my beloved husband, Bingzhen, who has quite simply been my strength and stay all these years. Without them, I would not be where I am today.

## Dedication

To my parents and husband, this thesis is dedicated to you.

# Table of Contents

**Examining Committee** **ii**

**Author's Declaration** **iii**

**Statement of Contributions** **iv**

**Abstract** **v**

**Acknowledgements** **vii**

**Dedication** **viii**

**List of Figures** **xiii**

**List of Tables** **xvi**

**1 Portfolio Selection – from under-diversification to concentration** **1**

1.1 Introduction . . . . . 1

1.2 Data and Statistical Analysis . . . . . 4

1.2.1 Data Description . . . . . 4

1.2.2 Statistical Analysis . . . . . 5

1.3 Empirical Results . . . . . 7

1.3.1 The Reconciliation of Concentration and Diversification . . . . . 7

1.3.2 Discussions . . . . . . . . . . 9
1.4 Conclusion . . . . . . . . . . 10

**2 A Functional Data Approach for Continuous-time Analysis Subject to Modeling Discrepancy under Infill Asymptotics 14**
2.1 Introduction . . . . . . . . . . 14
2.2 Methodology and Large Sample Properties . . . . . . . . . . 16
2.2.1 Fitting and Forecasting . . . . . . . . . . 16
2.2.2 Large Sample Properties . . . . . . . . . . 18
2.3 Simulation Analysis . . . . . . . . . . 21
2.3.1 The Data-generating Process . . . . . . . . . . 21
2.3.2 Fitting and Forecasting with FDA . . . . . . . . . . 22
2.3.3 Comparison to Parametric Methods . . . . . . . . . . 24
2.4 Conclusion . . . . . . . . . . 26

**3 Stochastic Discount Factors in Incomplete Markets 27**
3.1 Introduction . . . . . . . . . . 27
3.2 Discrete-time Setting . . . . . . . . . . 29
3.2.1 One Risk-free Bond, One Risky Asset, Two periods, Three States (1-1-2-3) . . . . . . . . . . 30
3.2.2 One Risk-free Bond, $A$ Risky Assets, Two Periods, $A+2$ States (1-$A$-2-$A+2$) . . . . . . . . . . 34
3.2.3 One Risk-free Bond, One Risky Asset, Three Periods, Three States (1-1-3-3) . . . . . . . . . . 38
3.2.4 Generalization of the Discrete-time Setting . . . . . . . . . . 41
3.3 Continuous-time Setting . . . . . . . . . . 44
3.3.1 Set Properties . . . . . . . . . . 46
3.3.2 Measure for Market Incompleteness . . . . . . . . . . 48
3.4 Application . . . . . . . . . . 49

3.4.1 Data Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
3.4.2 Estimation Algorithm . . . . . . . . . . . . . . . . . . . . . . . . . . 51
3.4.3 Estimation Results . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
3.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

**References** **57**

**APPENDICES** **69**

**A Appendices of Chapter 1** **70**
A.1 The Mean-Variance Optimization . . . . . . . . . . . . . . . . . . . . . 70
A.2 The DR Method . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
A.3 The EVT Method . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 72

**B Appendices of Chapter 2** **76**
B.1 Additional Notations and Lemmas. . . . . . . . . . . . . . . . . . . . . 76
B.2 Proof of Theorem 2.2.1 . . . . . . . . . . . . . . . . . . . . . . . . . . 77
B.3 Proof of Theorem 2.2.2 . . . . . . . . . . . . . . . . . . . . . . . . . . 78
B.4 Proof of Lemma 2.2.1 . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
B.5 Proof of Lemma B.1.1 . . . . . . . . . . . . . . . . . . . . . . . . . . . 80
B.6 Proof of Lemma B.1.2 . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
B.7 Proof of Lemma B.1.3 . . . . . . . . . . . . . . . . . . . . . . . . . . . 83
B.8 FDA Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 84
B.9 MLE Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 88

**C Appendices of Chapter 3** **95**
C.1 Proofs of Propositions and Theorems . . . . . . . . . . . . . . . . . . . 95
C.2 Proof of Proposition 3.2.1 . . . . . . . . . . . . . . . . . . . . . . . . . 96
C.3 Proof of Theorem 3.2.1 . . . . . . . . . . . . . . . . . . . . . . . . . . 96

C.4 Proof of Lemma 3.2.1 . . . . . . . . . . 99
C.5 Proof of Theorem 3.2.2 . . . . . . . . . . 99
C.6 Proof of Proposition 3.2.2 . . . . . . . . . . 101
C.7 Proof of Theorem 3.2.3 . . . . . . . . . . 102
C.8 Proof of Lemma 3.2.2 . . . . . . . . . . 102
C.9 Proof of Theorem 3.2.4 . . . . . . . . . . 103
C.10 Proof of Proposition 3.2.4 . . . . . . . . . . 104
C.11 Proof of Theorem 3.2.7 . . . . . . . . . . 105
C.12 Proof of Lemma 3.2.3 . . . . . . . . . . 106
C.13 Proof of Theorem 3.2.8 . . . . . . . . . . 107
C.14 Proof of Proposition 3.3.1 . . . . . . . . . . 108
C.15 Proof of Theorem 3.3.1 . . . . . . . . . . 109
C.16 Proof of Lemma 3.3.1 . . . . . . . . . . 109
C.17 Proof of Theorem 3.3.2 . . . . . . . . . . 109
C.18 Proof of Theorem 3.3.3 . . . . . . . . . . 111

# List of Figures

1.1 The Estimated Market Frontiers . . . . . . . . . . . . . . . . . . . . . . . . . . 8
1.2 Illustration of the True Market Frontier . . . . . . . . . . . . . . . . . . . . . . 10
1.3.1 The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), Brazil . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
1.3.2 The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), China . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
1.3.3 The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), India . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
1.3.4 The Estimated Market Frontiers (Panel A) and Its Left-lower orthant (Panel B), Japan . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
1.3.5 The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), UK . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
1.3.6 The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), US . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

2.1 Functional Data Prediction, Returns . . . . . . . . . . . . . . . . . . . . . . . 23
2.2 Functional Data Prediction, Volatility . . . . . . . . . . . . . . . . . . . . . . 24

3.1 Evolution of the Degree of Market Incompleteness . . . . . . . . . . . . . . . . . 55

A.1 Moment Estimator Versus $k$ for G1 of the China Market. . . . . . . . . . . . . . 74
A.2 Endpoint Estimators Versus $k$ for G1 of the China Market. . . . . . . . . . . . 75

B.1 Forecast vs. Underlying, $\hat{S}_D(t_j)$ . . . . . . . . . . . . . . . . . . . . . . 84

B.2 Forecast vs. Underlying, $\hat{S}_S(t_j)$ . . . . . . . . . . . . . . . . . . . . . . . . . . 85
B.3 Functional Data Prediction with an 8-month Rolling Window, Daily Returns 85
B.4 Functional Data Prediction with an 8-month Rolling Window, Daily Volatility 85
B.5 Forecast vs. Underlying, Daily $\hat{S}_S(t_j)$ . . . . . . . . . . . . . . . . . . . . . 86
B.6 Forecast vs. Underlying, Daily $\hat{S}_D(t_j)$ . . . . . . . . . . . . . . . . . . . . . 86
B.7 Functional Data Prediction with a 1-month Rolling Window, Continuous Returns . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 86
B.8 Functional Data Prediction with a 1-month Rolling Window, Continuous Volatility . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87
B.9 Forecast vs. Underlying, Continuous $\hat{S}_S(t_j)$ . . . . . . . . . . . . . . . . . . 87
B.10 Forecast vs. Underlying, Continuous $\hat{S}_D(t_j)$ . . . . . . . . . . . . . . . . . . 87
B.11 Functional Data Prediction with a 1-month Rolling Window, Daily Returns 88
B.12 Functional Data Prediction with a 1-month Rolling Window, Daily Volatility 88
B.13 Forecast vs. Underlying, Daily $\hat{S}_S(t_j)$ . . . . . . . . . . . . . . . . . . . . . 88
B.14 Forecast vs. Underlying, Daily $\hat{S}_D(t_j)$ . . . . . . . . . . . . . . . . . . . . . 89
B.15 Fitting DV Process Using a DV Model — $\hat{a}$ . . . . . . . . . . . . . . . . . . . 89
B.16 Fitting DV Process Using a DV Model — $\hat{\alpha}$ . . . . . . . . . . . . . . . . . . . 89
B.17 Fitting DV Process Using a DV Model — $\hat{\beta}$ . . . . . . . . . . . . . . . . . . . 90
B.18 Fitting SV Process Using a SV Model — $\hat{a}$ . . . . . . . . . . . . . . . . . . . 90
B.19 Fitting SV Process Using a SV Model — $\hat{\alpha}$ . . . . . . . . . . . . . . . . . . . 90
B.20 Fitting SV Process Using a SV Model — $\hat{\beta}$ . . . . . . . . . . . . . . . . . . . 90
B.21 Fitting SV Process Using a SV Model — $\hat{\sigma}$ . . . . . . . . . . . . . . . . . . . 91
B.22 Fitting SV Process Using a SV Model — $\hat{\rho}$ . . . . . . . . . . . . . . . . . . . 91
B.23 Fitting SV Process Using a DV Model — $\hat{a}$ . . . . . . . . . . . . . . . . . . . 91
B.24 Fitting SV Process Using a DV Model — $\hat{\alpha}$ . . . . . . . . . . . . . . . . . . . 91
B.25 Fitting SV Process Using a DV Model — $\hat{\beta}$ . . . . . . . . . . . . . . . . . . . 92
B.26 Fitting DV Process Using a SV Model — $\hat{a}$ . . . . . . . . . . . . . . . . . . . 92
B.27 Fitting DV Process Using a SV Model — $\hat{\alpha}$ . . . . . . . . . . . . . . . . . . . 92

B.28 Fitting DV Process Using a SV Model — $\hat{\beta}$ . . . . . . . . . . . . . . . . . . . . . . 92
B.29 Fitting DV Process Using a SV Model — $\hat{\sigma}$ . . . . . . . . . . . . . . . . . . . . . . 93
B.30 Fitting DV Process Using a SV Model — $\hat{\rho}$ . . . . . . . . . . . . . . . . . . . . . . 93
B.31 FDA vs. Misspecified MLE, Relative MSFE for SV Data . . . . . . . . . . . . . 93
B.32 FDA vs. Misspecified MLE, Relative MSFE for DV Data . . . . . . . . . . . . 94
B.33 FDA vs. Correctly Specified MLE, Relative MSFE for SV Data . . . . . . . . . 94
B.34 FDA vs. Correctly Specified MLE, Relative MSFE for DV Data . . . . . . . . 94

# List of Tables

1.1 Summary Statistics of Market Portfolios under Diversification . . . . . . . . . 6
1.2 EVT Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.1 Estimating GARCH Parameters with MLE . . . . . . . . . . . . . . . . . . . . . 25

# Chapter 1

# Portfolio Selection – from under-diversification to concentration[1]

## 1.1 Introduction

Many times, the two opposing investment strategies, diversification and concentration, have been directly compared (e.g., Bird et al., 2012; Yeung et al., 2012; Choi et al., 2017). As shown in the path-breaking work by Markowitz (1952), diversification has the strength of lowering overall portfolio risk, yet may also limit portfolio performance (e.g., Evans and Archer, 1968; Fisher and Lorie, 1970; Liu, 2016). On the flip side, concentration is known for its potential to generate better than market average returns, but is criticized for its lack of ability to avert idiosyncratic risks (e.g., Keynes et al., 1983; Buffett, 1994; Loeb, 2007; Ekholm and Maury, 2014).

While there is much less dispute adopting Markowitz's approach as the benchmark for diversification, the exact meaning of concentration in portfolio selection is still vague. For instance, studies have used the number of stocks in a portfolio to distinguish between diversification and concentration, and thresholds of 3, 11 to 15, or 30 have been proposed (e.g., Statman, 1987; Goetzmann and Kumar, 2008; Ivković et al., 2008). However, this measure is arbitrary, and there has not yet been a consensus on the threshold (Campbell, 2018, pp.324; Oyenubi, 2019). Another metric – the concentration index – estimates how

[1]This chapter is co-authored with Jiawen Xu, Kai Liu, and Tao Chen

much a portfolio deviates from the benchmark (Blume and Friend, 1975), with a lower index value representing a higher level of diversification and a zero index value representing "the diversified" portfolio (e.g., Brands et al., 2005; Kacperczyk et al., 2005; Goldman et al., 2016). Yet, this index only compares the relative concentration level of two portfolios, and "the concentrated" portfolio is either undefined or reduced to a single asset, which is not an interesting case. There is also the normalized portfolio variance, which is the ratio of portfolio variance over the average variance of stocks in that portfolio (e.g., Goetzmann et al., 2005; Goetzmann and Kumar, 2008). Assuming equal weights, a lower value of normalized variance suggests more diversification, and again this is a relative measure with "the concentrated" portfolio undefined.

As reminded by the motivating example in (Markowitz, 1952, pp.78-79), diversification prevails when there are multiple assets generating maximum expected return due to a simple law of large numbers argument. Therefore, the discussion of concentration is sensible only when confronting assets with heterogeneous expected returns, and it is evident that the three commonly used measures above do not extend to concentration in natural ways. In this paper, we decide not to dwell on the concept of under-diversification, which generalizes any deviations from diversification, but to propose a novel definition for concentration together with an estimation method for its empirical implementation.

Our inspiration comes from the example in (Campbell, 2018, pp.329-330), where one risky asset is only held by a group of homogeneous investors with identical initial wealth together with the riskless asset. When one-period power utility and log-normal payoff are imposed, this simple setup generates a crucial implication for the current paper: *undiversified investors care about the total variance of an asset's return and not just its covariance with a broader index, idiosyncratic volatility is positively related to expected return.*

Now we are ready to motivate our definition of concentration. We view that the market is dominated by undiversified investors, individuals, or institutions, who are constrained by their own investment policies and time horizons. They are heterogeneous in that they have different perceptions of the market, initial wealth, risk preferences, and inter-temporal substitution. Based on this set of beliefs, instead of assuming that investors are balancing the expected return and variance trade-off globally (as proposed by Markowitz and his followers), we assume that they choose or are only allowed to take a certain level of risk and strive for the highest possible expected return at that local variance. Hence, the determinant factor of investors' behavior is the standalone variances of assets' returns in the portfolio as opposed to their covariance in the diversification (Markowitz's) sense. Further, note that our definition of concentration is in terms of a fixed portfolio variance, which is assumed to be investor specific.

Aggregation of heterogeneous agents has always been a challenge in Macroeconomics. From the newly defined concentration's perspective, we are able to construct the frontier (hereafter, concentrated market frontier) in the following way: let portfolio variance sweep through its domain, and at each level, there is a projected maximum expected return associated with it. Then, the collection of those pairs forms the concentrated market frontier. To be conformable with Markowitz's framework, we replace the variance coordinate with its positive root. It is important to note that this concentrated market frontier should only be viewed as one extreme, similar to (or as unrealistic as) the well-known diversified market frontier[2] and it is actually the combination of both frontiers that provides us further insights.

The empirical implementations of the two strategies are summarized as follows, and we will elaborate in Section 1.2.2. The diversified market frontier is estimated through a dimensional reduction (DR) technique introduced by Liu (2017) and followed by the mean-variance optimization in Markowitz (1952) and Merton (1972); the concentrated market frontier is estimated by an extreme value theory (EVT)-based method originally proposed by Chen and Yang (2020) for general frontiers.

Combining the two estimated frontiers, we find (i) an implied risk threshold, where the two opposing investment strategies reconcile. When the risk level is below the threshold, diversification yields a higher expected return, whereas once the risk level exceeds the threshold, concentration becomes the dominant strategy. Different from the traditional concave efficient frontier, the merged frontier reveals to be seagull-shaped. This mixed frontier is piecewise concave and robust to all major stock markets we study. (ii) The family of frontiers nested inbetween the estimated curves captures the true market frontier, which stems from an unknown distribution of undiversified investors. From a set-identified perspective, we take the equity premium puzzle as an example to illustrate the potential of this collection of market portfolios to provide fresh insights into the related researches.

The paper proceeds as follows. The data and the statistical methods employed to estimate the frontiers are explained in Section 1.2. We present and discuss our findings in Section 1.3. Section 1.4 concludes.

[2]A concise way of saying the estimated efficient frontier implied by Markowitz's mean-variance optimization method.

## 1.2 Data and Statistical Analysis

We start with the data description and then the procedures to implement the two investment strategies.

### 1.2.1 Data Description

The data are collected monthly on six stock pools around the world, three of which are emerging markets, namely Brazil, China, and India, and the remaining three are developed markets, namely Japan, the United Kingdom (UK), and the United States (US)[3]. The three selected emerging markets are among the largest and fastest-growing emerging markets, while the three selected developed economies are among the most significant stock markets worldwide. Moreover, as each of the six markets has been extensively studied in the financial literature (e.g., Hamao et al., 1990; Floros, 2005; Gay Jr et al., 2008; Madaleno and Pinho, 2012; Fang and You, 2014), they serve as appropriate examples for demonstrating the robustness of our findings.

Our data is from January 2000 to December 2021 since 2000 is the first year for which we are able to obtain complete stock data for all six stock markets. The monthly logarithmic return is used, and the monthly volatility is the positive root derived from its variance (Andersen and Bollerslev, 1998; Feibel, 2003). Both the expected return and volatility are annualized according to Jorion (2010, pp.224). To obtain relatively stable portfolio variance, stocks with less than four years of return data are excluded from this analysis. We then sort our sample in ascending order by volatility and trim the boundary observations by excluding the first and last 5% of the data[4]. The sample sizes of the four markets (in the aforementioned order) are 450, 3240, 3930, 3060, 2340, and 4710, respectively.

[3]The Brazil stock pool includes all stocks currently listed at the B3 Stock Exchange. The China stock pool includes all stocks currently listed at the Shanghai and the Shenzhen Stock Exchanges. The India stock pool includes all stocks currently listed at the Bombay Stock Exchange and the National Stock Exchange of India. The Japanese stock pool includes all stocks currently listed at the Tokyo Stock Exchange. The US stock pool includes all stocks currently listed at the New York Stock Exchange and the NASDAQ Stock Market. The UK stock pool includes all stocks currently listed at the London Stock Exchange.

[4]The method to estimate the concentrated market frontier is non-parametric, which is generally known for boundary bias. This is the reason we follow the common practice in non-parametric estimation by trimming the observations close to the boundaries. To ensure comparability, both frontiers are estimated based on the same trimmed data set.

### 1.2.2 Statistical Analysis

This section explains the statistical methods utilized to estimate the diversified and concentrated market frontiers. Both procedures are introduced to the literature recently. With no intention we try to claim that they are "the" procedures associated with both frontiers; this paper employs them as viable estimation tools.

#### The Diversified Market Frontier

Ideally, the diversified market frontier can be directly estimated following the mean-variance optimization described in Markowitz (1952). To improve readability, we move the mathematical formulation to Appendix A.1. Though conceptually intuitive, Markowitz's optimization method is challenging to implement empirically: the sample covariance matrix is known to suffer from singularity or near-singularity problems in the presence of a large number of assets (Buser, 1977; Pappas et al., 2010). In order to restore invertibility of the covariance matrix in certain sense, shrinkage method of different kinds have been proposed (e.g., Leonard et al., 1992; Yang and Berger, 1994; Daniels and Kass, 1999; Ledoit and Wolf, 2003, 2004, 2012, 2017); however, all these methods rely heavily on the interpretation of priors or the Bayesian likelihood-based foundation.

We resort to a pure numerical algorithm that was initially proposed by Liu (2017). The main reason behind this choice is because their DR approach does not require many technical assumptions. It simply picks out $N_{DR}$ number of assets with the highest Sharpe Ratios that contribute to a pre-specified percentage of the total market variation and can be arbitrarily close to the unit. This paper works with 95%. With these chosen stocks, we then solve the mean-variance optimization to obtain the diversified market frontier. For the sake of completeness, the mathematical derivations can be found in Appendix A.2.

Each market portfolio is derived using the capital market line, where the corresponding average one-year treasury rate from 2000 to 2021 is used as proxy for the risk-free interest rate, and we summarize the statistics in Table 1.1. The three emerging markets appear to have considerably greater volatility than the developed ones, which accords the literature that emerging markets are inherently riskier (e.g., Sharkasi et al., 2006; Saranya and Prasanna, 2014). The Sharpe ratio shows that the US market portfolio has the best risk-adjusted performance among the six markets, implying that it compensates the investor for the risk taken most effectively. Furthermore, the sample sizes are all significantly smaller than the original ones, which suggests that a substantial number of stocks are highly correlated in their returns.

Table 1.1: Summary Statistics of Market Portfolios under Diversification

| | Mean | Standard Deviation | Sharpe Ratio | Number of Assets |
|---|---|---|---|---|
| Brazil | 0.365 | 0.208 | 1.163 | 24 |
| China | 0.440 | 0.341 | 1.215 | 34 |
| India | 0.440 | 0.320 | 1.166 | 38 |
| Japan | 0.185 | 0.113 | 1.629 | 57 |
| UK | 0.375 | 0.145 | 2.441 | 58 |
| US | 0.264 | 0.079 | 3.118 | 127 |

**The Concentrated Market Frontier**

In Finance, EVT has been successfully employed to estimate the extreme quantiles, tail probabilities of asset returns and risks (e.g., Pownall and Koedijk, 1999; Longin, 2000; McNeil and Frey, 2000; Rocco, 2014; Longin, 2016). Here we briefly describe the idea of the EVT-based method proposed by Chen and Yang (2020) and leave the details to Appendix A.3.

If we view uncertainty and the expected return as the input and output of a production function, then predicting the maximum possible output based on similar inputs matches exactly the way we construct the concentrated frontier. Specifically, we divide the trimmed sample into $g$ equal-sized groups, and this analysis takes $g$ as 10 (e.g., for the China market, each group consists of 324 observations). The choice of $g$ follows the fundamental bias-variance compromise in that if $g$ gets bigger, the observations within each group are more "similar" but with less number of them. In this case, the bias reduces while the variance of the estimator increases, and a reversed argument holds when $g$ becomes smaller. For each group (G1-G10) in the stock pool, we first calculate the moment estimator of the extreme-value index, $\hat{\gamma}$, according to Dekkers et al. (1989). Subsequently, this $\hat{\gamma}$ is used to estimate the maximum expected return under EVT (hereafter, EVT expected return) at each risk level based on von Mises (1954) and Jenkinson (1955).

Table 1.2 presents the EVT and the maximum expected returns for each group. The relatively large $\hat{\gamma}$'s for the groups with greater risks indicate that considerable improvement in investment performance can be attained at higher risk levels by concentration. In contrast, little or no increase in return can be achieved at the lower risk levels with relatively small $\hat{\gamma}$'s. Last, we connect all pairs of EVT expected return and volatility to obtain the concentrated market frontier.

Table 1.2: EVT Results

| | | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | G10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Brazil | $\hat{\gamma}$ | -0.272 | -0.401 | -0.334 | -0.217 | -0.216 | -0.130 | -0.144 | -0.228 | -0.245 | -0.282 |
| | EVT Return | 0.334 | 0.392 | 0.441 | 0.489 | 0.526 | 0.568 | 0.610 | 0.649 | 0.762 | 0.923 |
| | Max Return | 0.333 | 0.334 | 0.345 | 0.413 | 0.417 | 0.257 | 0.279 | 0.648 | 0.562 | 0.858 |
| China | $\hat{\gamma}$ | -0.326 | -0.253 | -0.207 | -0.167 | -0.142 | -0.122 | -0.110 | -0.114 | -0.246 | -0.168 |
| | EVT Return | 0.312 | 0.423 | 0.500 | 0.559 | 0.594 | 0.625 | 0.648 | 0.677 | 0.720 | 0.796 |
| | Max Return | 0.311 | 0.410 | 0.335 | 0.359 | 0.538 | 0.436 | 0.597 | 0.503 | 0.719 | 0.679 |
| India | $\hat{\gamma}$ | -0.262 | -0.181 | -0.160 | -0.160 | -0.122 | -0.156 | -0.196 | -0.181 | -0.101 | -0.093 |
| | EVT Return | 0.509 | 0.592 | 0.653 | 0.707 | 0.751 | 0.788 | 0.820 | 0.854 | 0.887 | 0.955 |
| | Max Return | 0.508 | 0.431 | 0.546 | 0.643 | 0.571 | 0.628 | 0.721 | 0.731 | 0.536 | 0.929 |
| Japan | $\hat{\gamma}$ | -0.167 | -0.167 | -0.117 | -0.167 | -0.158 | -0.151 | -0.150 | -0.546 | -0.230 | -0.168 |
| | EVT Return | 0.186 | 0.249 | 0.288 | 0.325 | 0.364 | 0.407 | 0.460 | 0.515 | 0.563 | 0.602 |
| | Max Return | 0.185 | 0.229 | 0.170 | 0.249 | 0.319 | 0.307 | 0.317 | 0.514 | 0.506 | 0.568 |
| UK | $\hat{\gamma}$ | -0.288 | -0.257 | -0.194 | -0.231 | -0.156 | -0.248 | -0.153 | -0.233 | -0.316 | -0.534 |
| | EVT Return | 0.351 | 0.483 | 0.575 | 0.656 | 0.752 | 0.839 | 0.924 | 1.030 | 1.137 | 1.333 |
| | Max Return | 0.282 | 0.408 | 0.393 | 0.482 | 0.512 | 0.709 | 0.626 | 0.880 | 1.037 | 1.333 |
| US | $\hat{\gamma}$ | -0.255 | -0.133 | -0.107 | -0.105 | -0.121 | -0.089 | -0.068 | -0.078 | -0.094 | -0.080 |
| | EVT Return | 0.254 | 0.447 | 0.635 | 0.783 | 0.897 | 0.985 | 1.072 | 1.159 | 1.226 | 1.316 |
| | Max Return | 0.254 | 0.309 | 0.428 | 0.426 | 0.541 | 0.699 | 0.394 | 0.610 | 0.565 | 0.488 |

## 1.3 Empirical Results

We overlay the estimated diversified and concentrated market frontiers in Figure 1.1 to reconcile the two investment strategies. Section 1.3.2 provides further Economic insights through market portfolio.

### 1.3.1 The Reconciliation of Concentration and Diversification

When combining the two market frontiers, we observe a seagull-shaped frontier in all markets. Figure 1.1 shows that there exists an implied risk threshold, at which the two strategies reconcile. When the risk is below that threshold, diversification yields higher expected returns, while concentration becomes superior when the risk exceeds that threshold. This finding is consistent with the existing literature on the trade-off between expected return and variance (e.g., Modigliani and Leah, 1997; Campbell et al., 2001; Ghysels et al.,

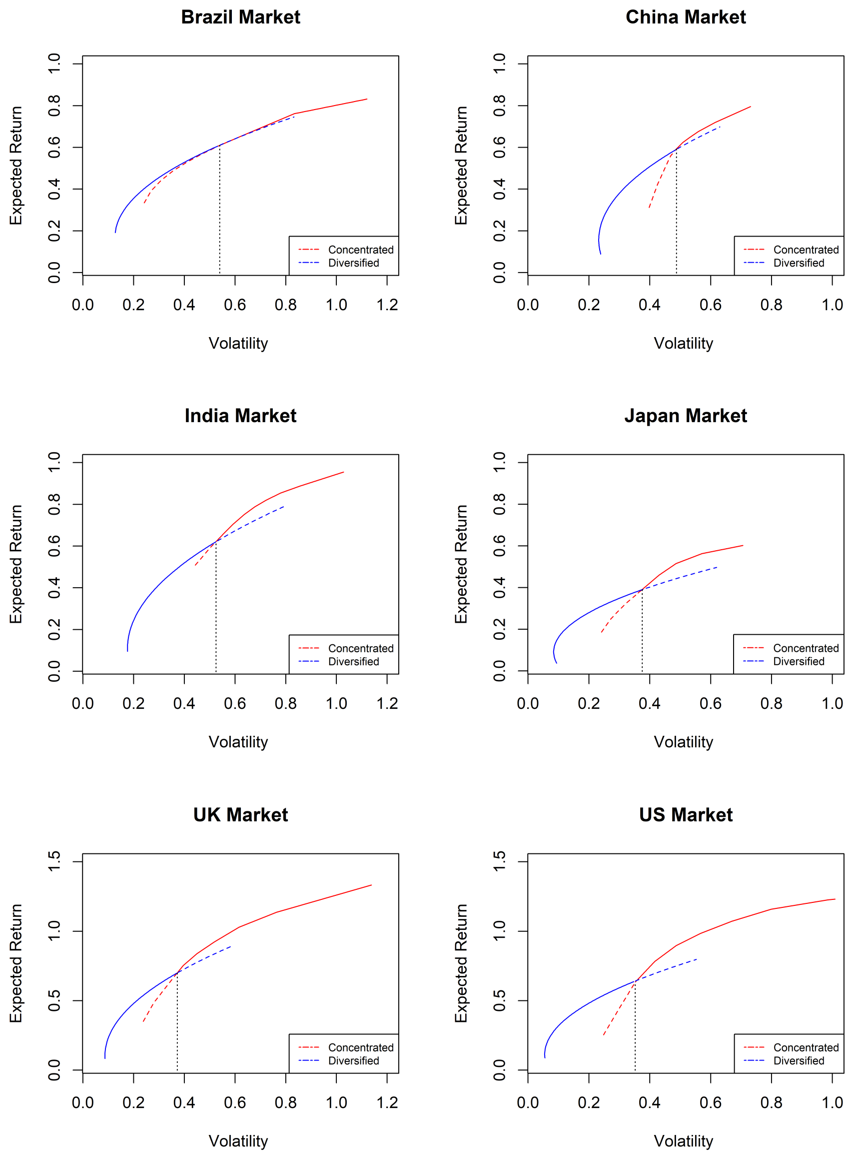


Figure 1.1: The Estimated Market Frontiers

2005; Bali and Peng, 2006). In particular, if the investor is constrained at a relatively low-risk level, diversification outperforms concentration by reducing the potential idiosyncratic risk of the portfolio without sacrificing the potential return much. While concentration brings significant risk, practitioners such as Warren Buffett advocate, it can also result in a significant gain if the investor makes the proper judgment. Additionally, in line with our discussion in Table 1.1, concentration outperforms diversification at the lowest risk threshold for the US market, whence risk-taking activities are compensated most effectively.

### 1.3.2 Discussions

Here we explore possible Economic insights from those two estimated frontiers, which should only be viewed as the two extreme cases. Between them, there is a family of frontiers representing under-diversification of different degrees. Figure 1.2 illustrates this idea through two sample intermediary curves. If we use the diversified market frontier as the benchmark and move the concentrated one to the left while rotating clock-wisely, we are able to trace out all the possible frontiers representing an under-diversified investment strategy with an increasing degree of diversification till it meets the fully diversified frontier. Within this process, the curvature of frontiers changes.

Figure 1.3 presents each frontier paired with its corresponding capital market line, and the tangent point $M$ is interpreted as the market portfolio. We use the subscripts $d$ and $c$ to differentiate the market portfolio associated with the diversified and concentrated frontiers. Most of the time, researches use general stock indexes as proxies for $M_d$. Assuming agents are allowed to borrow and lend freely at the risk-free rate $R$, portfolio $M_c'$ is introduced as a device to replicate the true market portfolio that matches the actual data, which shares the same Sharpe ratio as $M_c$ and admits the same level of risk as $M_d$. Panel A's in Figure 3 present the geometric relationships with Panel B being the left-lower orthants respectively.

The distance between $M_d$ and $R$ is the well-known Equity Premium (EP) proposed by Mehra and Prescott (1985), which has been extensively discussed over the past 35 years (Mehra, 2008). By Adopting the same consumption aggregation idea in Caselli and Ventura (2000), we conduct the analysis assuming that the sum of heterogeneous investors behave as if it were a single agent. Then, EP should be measured as the distance between $M_c'$ and $R$, which can be substantially smaller than the traditional one. Here it is important to note that we are not intending to add another trial to tackle the EP puzzle. Our goal is to use it as an example to show the practicality of our findings for any research involving the market portfolio.

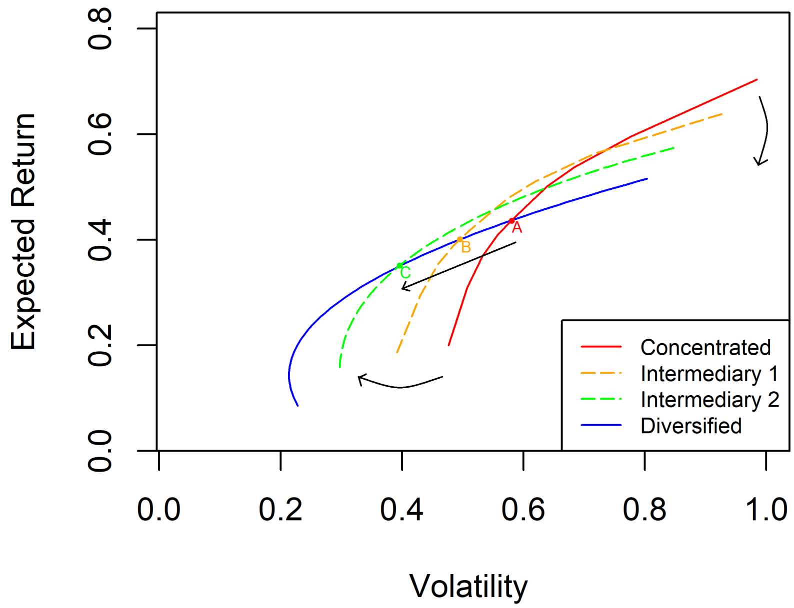


Figure 1.2: Illustration of the True Market Frontier

## 1.4 Conclusion

This study proposes a novel definition of concentration for portfolio selection and creates a new market frontier based on a well-known aggregation method in Macroeconomics. A comparison to the traditional diversified portfolio selection approach is presented, and our estimated seagull-shaped frontier can be used as a guide when developing investment strategies so that, to maximize returns, investors can construct their portfolios either by diversification or concentration according to their risk constraints relative to the implied risk threshold. Further, possible implications of the new market portfolio are discussed, and it will be of compelling interest to locate a frontier from the family of frontiers we proposed that matches the data, and we defer this to further studies.

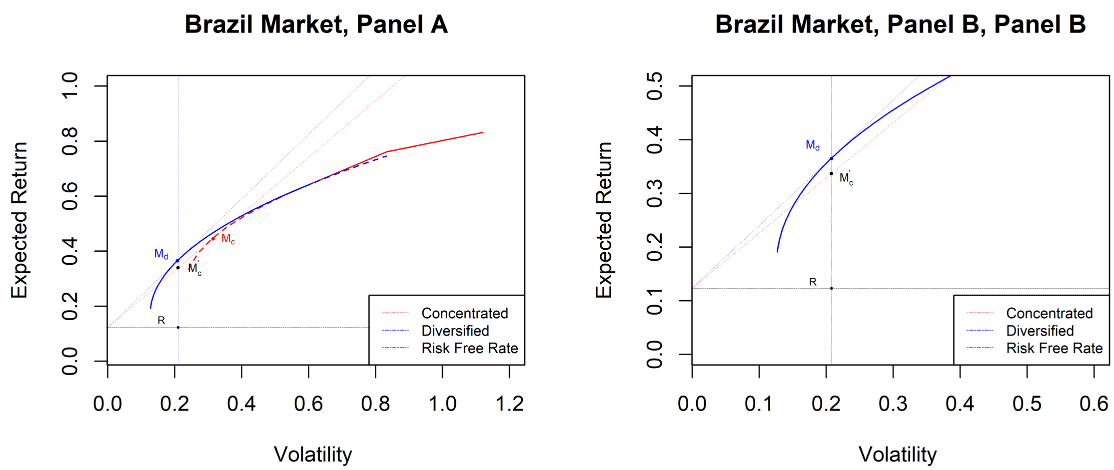


Figure 1.3.1: The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), Brazil

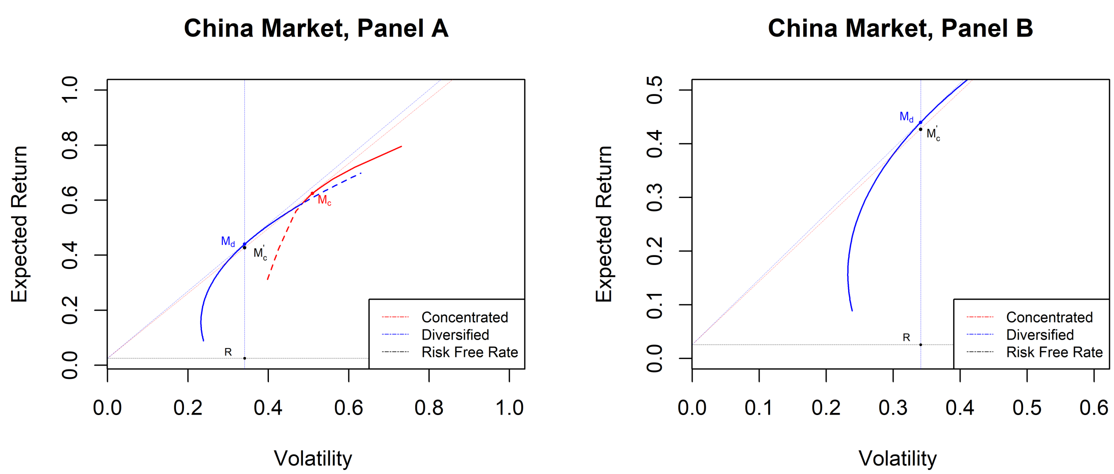


Figure 1.3.2: The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), China

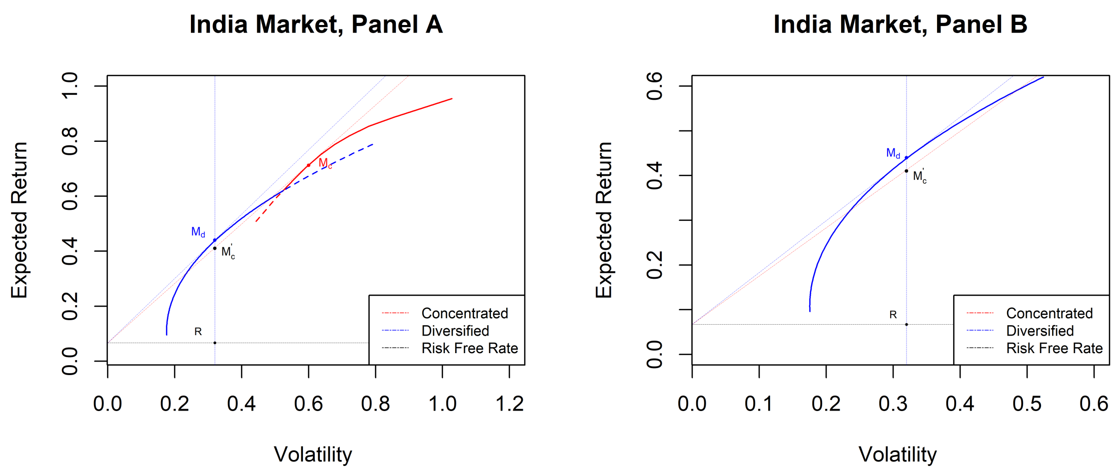


Figure 1.3.3: The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), India

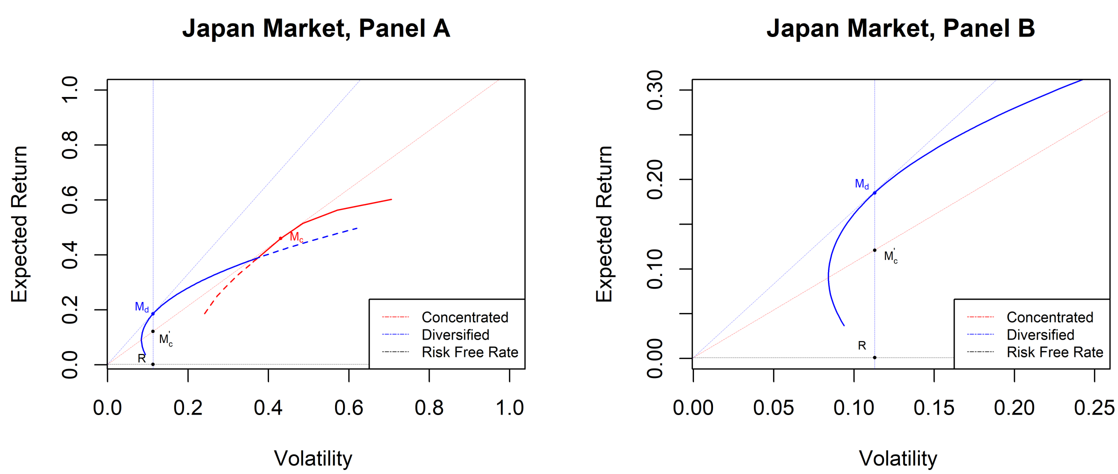


Figure 1.3.4: The Estimated Market Frontiers (Panel A) and Its Left-lower orthant (Panel B), Japan

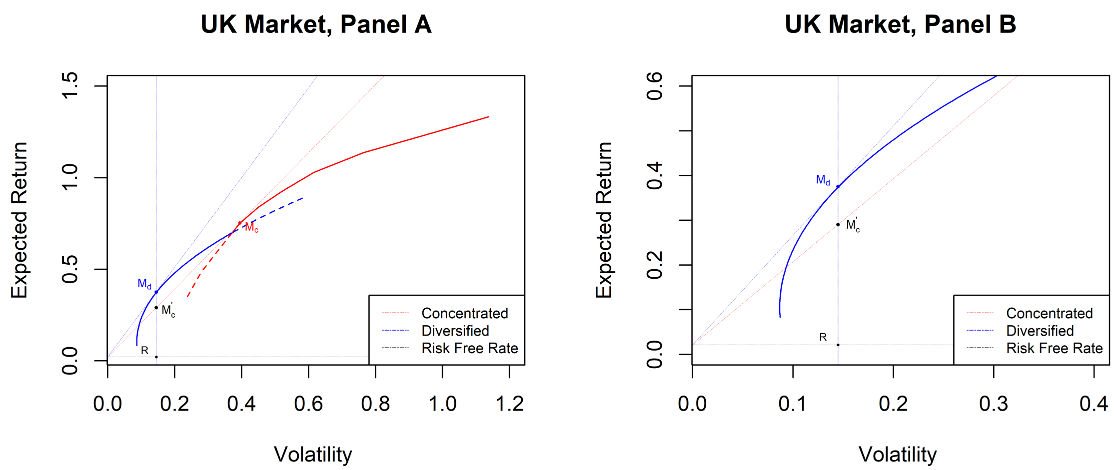


Figure 1.3.5: The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), UK

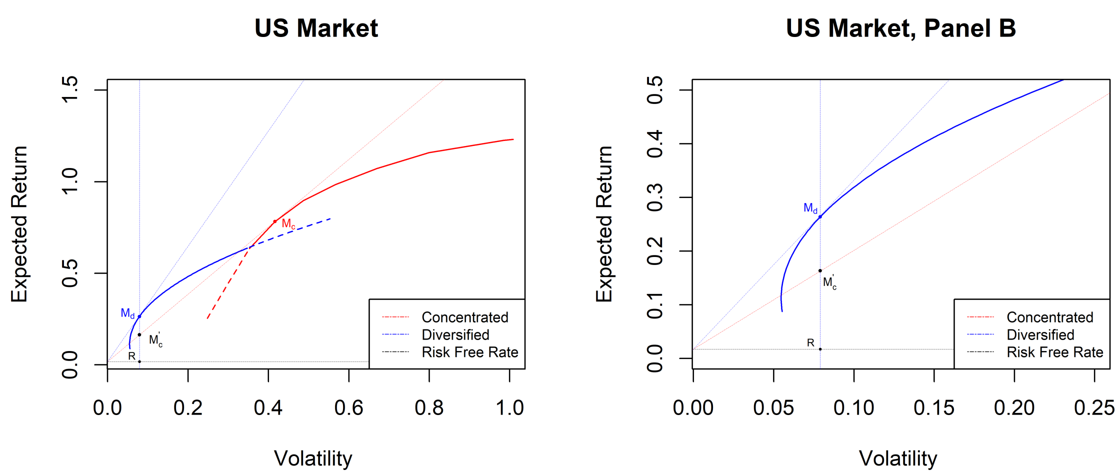


Figure 1.3.6: The Estimated Market Frontiers (Panel A) and Its Left-lower Orthant (Panel B), US

# Chapter 2

# A Functional Data Approach for Continuous-time Analysis Subject to Modeling Discrepancy under Infill Asymptotics[1]

## 2.1 Introduction

Much of finance and economics is about the study of dynamics over time, where analysis using time series data plays a vital part. Despite discrete-time observations in practice, many time series data such as stock prices, interest rates, and GDP are essentially drawn from their continuous-time underlying processes, whence consistent estimation can be achieved from continuous-time analyses but not necessarily from their discrete-time counterparts. Thus, the former has become increasingly incorporated by modern time series analysis (e.g., Merton, 1980; Merton and Samuelson, 1992; Stentoft, 2011; Buccheri et al., 2021).

Traditionally, when a continuous-time analysis is conducted parametrically, the modeling routinely originates in a discrete-time setting and is then generalized to a continuous-time formulation as the limit (Merton and Samuelson, 1992). Yet, the parametric generalization often places strong restrictions on the specifications of the convergence rates of frequency-dependent (hereafter, f.d.) parameters, resulting in models that are unadaptable to real-world data, and thus forecasting failure (Corradi, 2000; Das, 2002; Duan et al.,

[1]This chapter is co-authored with Renfang Tian and Tao Chen

2006; Trifi, 2006; Giraitis et al., 2007; Stentoft, 2011; Badescu et al., 2017; Hafner et al., 2017). For instance, with different sets of conditions regarding the speed of convergence of parameters, many GARCH-like processes have various sorts of diffusion processes (e.g., Nelson, 1990; Corradi, 2000; Duan et al., 2006; Hafner et al., 2017). Consequently, there arises discrepancy in the continuous-time generalization as it is debatable which assumption is the correct one (e.g., Singleton, 2009, pp.176-178; Alexander and Lazar, 2021), and mistaking the assumption will lead to an invalid limit, thereby unreliable analysis results. In such circumstances, nonparametrics appear to be an appealing tool that can naturally adapt to the true limits based on data-driven methods and bypass this discrepancy in continuous-time modeling and analysis (Sundaresan, 2000).

There has been a large literature demonstrating the effectiveness of nonparametric methods for accurate estimation and forecasting in cases where parametric assumptions are deemed to be inadequate (e.g., Härdle et al., 1997; Heiler, 1999; Fan and Yao, 2003; Pena et al., 2011; Zhu et al., 2011; Bosq, 2012; Kutoyants, 2012; Kleppe et al., 2014; Ryabko, 2019; Aydin and Yilmaz, 2021). Many studies incorporated nonparametric methods to approximate the density of the states in the absence of a closed-form expression, so that they could use maximum likelihood to estimate continuous-time diffusion models (e.g., Durham and Gallant, 2002; Aït-Sahalia and Kimmel, 2007; Kleppe et al., 2014). Yet, they are still presuming the parametric format of the underlying processes. Another stream of literature used kernel-based methods to estimate the processes that are nonanticipative smooth functions with unknown structures (e.g., Kutoyants, 2012) and forecast through conditional density estimation (e.g., Härdle et al., 1997; Matzner-Løfber et al., 1998; Fan and Yao, 2003; Berry et al., 2015). It has been shown, however, that traditional kernel estimators are inconsistent under infill asymptotics over bounded domains (Lahiri, 1996; Bosq, 2012), and the use of kernel estimators under such an infill asymptotic structure requires careful consideration of the dependence among observations before large sample properties can be obtained, which requires substantial work (Kurisu, 2019).

The main contribution of our paper is to propose a fitting and forecasting approach from the viewpoint of functional data analysis (FDA) to accommodate f.d. data structures under infill asymptotics. In particular, we fit the underlying process by using local polynomials, and we obtain the forecasts by extending the movement of the process based on the boundary derivatives of the functional fitting. The FDA method has been widely employed in capturing the dynamics of unknown smooth functions behind stochastic processes (Ramsay, 2005; Zhu et al., 2011; Chen et al., 2018), but not specifically for the purpose of tracking the f.d. data structures in continuous-time analysis. This paper also makes a note that functional approaches can be further developed in this context to achieve desired performance under infill asymptotics with bounded or unbounded domains. The procedure and properties of the proposed method are illustrated through a simulation study. To improve readability, all proofs are placed in Appendix B.

The rest of this paper is structured as follows. In Section 2.2, we introduce the FDA-based method of fitting and forecasting continuous-time underlying processes and subsequently derive the large sample properties of the functional estimators. In Section 2.3, we take the strong GARCH(1,1) process as a motivating example and show via simulation that in continuous-time analysis, while parametric methods, such as maximum likelihood estimation (MLE), encounter discrepancy, the proposed FDA-based method can provide reliable estimation and forecast. Section 2.4 concludes.

## 2.2 Methodology and Large Sample Properties

In this section, we explain the functional fitting and forecasting procedures followed by a discussion on their large sample properties. As noted by Jouzdani and Panaretos (2021), smoothness assumptions are often imposed in FDA to accommodate incomplete and noisy observations, but those assumptions are not applicable for a large group of stochastic processes such as diffusion processes; on the other hand, fully observed sample paths can be assumed to incorporate continuous but non-differentiable underlying processes, yet continuous-time underlying processes are mostly not fully observed in practice. Hence, in this study, we fit functional data through the Bernstein polynomial approximation, which can adapt to continuous but not necessarily differentiable stochastic processes, and no "complete observation" assumption needs to be made.

### 2.2.1 Fitting and Forecasting

Consider a complete probability space $(\Omega, \mathcal{F}, P)$ and a re-scaled bounded time interval $[0,1]$. We denote a sample path for any given state $\omega \in \Omega$ over $[0,1]$ by the mapping $X_\omega(t) : [0,1] \to \mathbb{R}$. Realizations are drawn at countably many time points, as such

$$X_{\omega,t} = X_\omega(t) + \epsilon_{\omega,t}, \tag{2.1}$$

where $\epsilon_{\omega,t}$ is the observational error. At any given time point $t$, we denote a $\Delta$-step-ahead predictor by $\hat{X}_\omega(t + \Delta)$. If the sample path $X_\omega(t)$ is continuous, it has an arbitrarily close polynomial approximation according to the Weierstrass approximation theorem, and Bernstein polynomials $B_K(t, X_\omega)$ provide a specific format of polynomial approximation, as such for each $K \in \mathbb{N}$,

$$B_K(t, X_\omega) := \sum_{k=0}^{K} X_\omega \left(\frac{k}{K}\right) \binom{K}{k} t^k (1-t)^{K-k}. \tag{2.2}$$

Then we have the following lemma.

**Lemma 2.2.1.** *Consider $B_K(t, X_\omega)$ as in Equation (2.2) with the continuous sample path $X_\omega$ for all $\omega \in \Omega$. For any $\varepsilon > 0$, there exists some $\delta > 0$ that induces $|t_1 - t_2| \le \delta \implies |X_\omega(t_1) - X_\omega(t_2)| \le \varepsilon/2$ for all $t_1, t_2 \in [0,1]$, such that for all $K \ge \sup_{t\in[0,1]} |X_\omega(t)|/(2\delta^2\varepsilon)$,*

$$\sup_{t\in[0,1]} |B_K(t, X_\omega) - X_\omega(t)| \le \varepsilon.$$

Lemma 2.2.1 suggests that, fixing everything else, for any given $\varepsilon$, one needs a high-degree polynomial for approximation if the continuity of $X_\omega$ requires a very small $\delta$, while low-degree polynomial can be used to achieve desired approximation if a large $\delta$ suffices the given $\varepsilon$. This lemma also implies that while fitting any continuous process, the typical smoothness assumptions for functional data can be imposed on the Bernstein approximating polynomials $B_K(t, X_\omega)$, and the remaining estimation error can be controlled by the selection of the degree $K$, which takes into account the level of volatility in the underlying process $X_\omega$.

In the spirit of FDA (Ramsay, 2005), one can construct a B-spline representation for the polynomial approximation $B_K(t, X_\omega)$ as such

$$B_K(t, X_\omega) = \mathbf{\Phi}^\top(\cdot)\boldsymbol{C},$$

where $\mathbf{\Phi}$ is a B-spline basis that contains $Q$ basis functions, $\boldsymbol{C}$ is a vector of $Q$ corresponding basis coefficients, and the superscript "$\top$" indicates the transpose operation. Then equation (2.1) can be rewritten as

$$X_{\omega,t} \approx \mathbf{\Phi}^\top(t)\boldsymbol{C} + \epsilon_{\omega,t}. \tag{2.3}$$

With the predetermined basis functions and the observations $\{X_{t_j}\}_{j=1}^J$ at $\{t_j\}_{j=1}^J \subseteq [0,1]$, the coefficients $\boldsymbol{C}$ can be estimated by minimizing the penalized least squares $J^{-1}\sum_{j=1}^J [X_{t_j} - X(t_j)]^2 + \lambda \int_0^1 X^{(2)}(t)dt$ with a tuning parameter $\lambda$, which yields

$$\tilde{\boldsymbol{C}}_\omega := \left[\frac{1}{J}\sum_{j=1}^J \mathbf{\Phi}(t_j)\mathbf{\Phi}^\top(t_j) + \lambda \int_0^1 \mathbf{\Phi}^{(2)}(t)\left\{\mathbf{\Phi}^{(2)}(t)\right\}^\top dt\right]^{-1} \frac{1}{J}\sum_{j=1}^J \mathbf{\Phi}(t_j) X_{\omega,t_j}, \tag{2.4}$$

where the superscript "(2)" indicates the 2nd derivative of a function. In practice, $Q$ and $\lambda$ can be selected through some cross-validation algorithms as explained in Ramsay (2005). Henceforth, an estimator for $X_\omega(t)$ can be defined as

$$\tilde{X}_\omega(t) := \mathbf{\Phi}^\top(t)\tilde{\boldsymbol{C}}_\omega, \ \forall t. \tag{2.5}$$

Throughout the paper, we use $\tilde{X}_\omega$ and $\hat{X}_\omega$ to distinguish between the fitting and forecast results.

With the fitted process $\tilde{X}_\omega(t)$, the $\Delta$-step-ahead predictor $\hat{X}_\omega(t+\Delta)$ can be obtained by a truncated Taylor expansion with the first $R$ derivatives of the fitted functions, such that

$$\hat{X}_\omega(t+\Delta) := \sum_{r=0}^{R} \frac{1}{r!}\Delta^r \tilde{X}_\omega^{(r)}(t) = \sum_{r=0}^{R} \frac{1}{r!}(\Delta)^r \tilde{\boldsymbol{C}}_\omega^\top \mathbf{\Phi}^{(r)}(t). \tag{2.6}$$

The validity of the above Taylor expansion requires the fitted function $\tilde{X}_\omega(t)$ to be at least $R$ times continuously differentiable, while the underlying process $X_\omega(t)$ only need to be continuous. Due to the local polynomial property of the B-spline basis, one can adopt the B-spline basis of order $R+2$ to ensure the desired differentiability of $\tilde{X}_\omega(t)$.

### 2.2.2 Large Sample Properties

We now discuss the properties of our functional predictors based on the following Assumptions.

**Assumption 2.2.1.** *Consider a continuous sample path $X_\omega(t)$ as such for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|t_1 - t_2| \le \delta$ implies $|X_\omega(t_1) - X_\omega(t_2)| \le \varepsilon/2$. Then we have the following.*

- *(a) The sample path $X_\omega(t)$ is observed on a set of evenly-spaced time points.*
- *(b) The observational error $\epsilon_{\omega,t_j}$ is uncorrelated across $\{t_j\}_{j=1}^J \subseteq [0,1]$, with $\mathbb{E}[\epsilon_{\omega,t} \mid X_\omega(s), s < t] = 0$ and $Var[\epsilon_{\omega,t}] < c < \infty$ for all $t \in [0,1]$ and some constant $c$.*
- *(c) $Q \sim J^{\alpha_1}$ and $\lambda \sim J^{\alpha_2}$ with $0 < \alpha_1 < 1$ and $\alpha_2 < 0$.*

Assumption 2.2.1 (a) is a sufficient but not necessary condition. When the sampling frequency increases properly, for example as stated in Claeskens et al. (2009), the desired properties of the functional estimator can be achieved without requiring evenly-spaced time series observations. A relaxed version is to assume that the ratio between the lengths of the largest and the smallest time windows is bounded away from zero and bounded above. For simplicity, we assume equal-spaced observations without loss of generality. Part (b) states a sufficient condition to achieve the consistency of the functional estimator $\tilde{X}_\omega$ as in Equation (2.5), while a proper "low correlation" assumption for this error term is also sufficient. Finally, Part (c) indicates that the dimension of the basis expansion shall increase with the sample size to allow for a consistent functional estimator $\tilde{X}_\omega$, and meanwhile, the tuning parameter shall be kept as a $o_p(1)$ so that the estimation error introduced by the roughness penalty dies down as the sample size grows towards infinity.

### Convergence in linear functionals

For any given $\Delta$, we consider the convergence of the forecasting process $\hat{X}_\omega$ over $(\Delta, 1)$, or equivalently, over $t \in (0, 1-\Delta)$. Under Assumption 2.2.1, we have the following results.

**Theorem 2.2.1** (Weak Convergence of $\hat{X}_\omega$ given $\Delta$)**.** *Consider a continuous sample path $X_\omega$ on $[0,1]$ as such for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|t_1 - t_2| \le \delta$ implies $|X_\omega(t_1) - X_\omega(t_2)| \le \varepsilon/2$.*

*Let $\mathcal{F}_R$ be a collection of all degree-R polynomials over $[0, 1-\Delta]$ for $R \ge K$, and $F_K: \mathcal{F}_R \to \mathbb{R}$ be the linear functional as such $F_K(\psi) := \left\langle \hat{X}_\omega(\cdot + \Delta) - B_K(\cdot + \Delta, X_\omega), \psi \right\rangle$ for $\psi \in \mathcal{F}_R$.*

*If Assumption 2.2.1 holds and given $K$, then $F_K(\psi) \xrightarrow{p} 0$ as $J \to \infty$ for all $\psi \in \mathcal{F}_R$.*

*If Assumption 2.2.1 holds and $K \to \infty$, then $\sup_{t\in(0,1-\Delta)} |\hat{X}_\omega(t+\Delta) - X_\omega(t+\Delta)| \xrightarrow{p} 0$ as $J \to \infty$.*

Theorem 2.2.1 establishes the convergence of the forecasting process $\hat{X}_\omega$ through its weak convergence to the Bernstein approximating polynomial $B_K(t+\Delta, X_\omega)$ and the uniform convergence of $B_K(t, X_\omega)$ to $X_\omega$ as $K$ increases towards infinity. Basically, $\hat{X}_\omega$ converges uniformly to the true path $X_\omega$ with enlarging $K$ and $J$; while with a fixed $K$, $\hat{X}_\omega$ still converges to the approximating polynomial $B_K(t+\Delta, X_\omega)$. Note that weak convergence implies uniform convergence over intervals on the real line, but the converse is not generally true.

### Asymptotic normality

Now we explore the asymptotic distribution of the functional prediction $\hat{X}_\omega$. However, before we state any further assumptions or the resulting asymptotic properties, it is important to note that the behavior of the observational error $\epsilon_{\omega,t}$ may affect the perspective of the realizations from $X_\omega$ and can also destabilize the asymptotic distribution of $\hat{X}_\omega$. For example, when a continuous sample path $X_\omega$ is observed without noise, it is more plausible to consider the functional fitting as an interpolation – in this case, we do not necessarily have asymptotic normality for $\hat{X}_\omega$. Yet, if $X_\omega$ is observed with noise, then we can establish a pointwise asymptotic normality under sufficient conditions.

**Assumption 2.2.2.** *(a) The noise $\epsilon_{t_j}$ is independent across all $\{t_j\}_{j=1}^J \subseteq [0,1]$.*

*(b) $Q \sim J^{\alpha_1}$ and $\lambda \sim J^{\alpha_2}$ with $\alpha_1 > 0$ and $\alpha_2 < 0$, such that $Q\lambda = o_p(J^{-1/2})$.*

Assumption 2.2.2 (a) states a sufficient condition for generating the asymptotic normality by the Lyapunov CLT which allows for unidentically distributed variables. If one has mixing random processes, a different version of this assumption can be specified by imposing identical distribution, where a different CLT can be applied without affecting the result on asymptotic normality. Part (b) assumes a stronger condition on the orders of the two parameters $Q$ and $\lambda$, so that the non-normal part of the estimation error is dominated and will not be inflated while deriving the asymptotic distribution.

**Theorem 2.2.2.** *Consider a continuous sample path $X_\omega$ on $[0,1]$ as such for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|t_1 - t_2| \leq \delta$ implies $|X_\omega(t_1) - X_\omega(t_2)| \leq \varepsilon/2$.*

*Under Assumptions 2.2.1 and 2.2.2, the following asymptotic normality holds for all $t \in (0, 1-\Delta)$ and $R \to \infty$ as $J \to \infty$:*

$$[V(t;\Delta,R,\mathbf{\Phi})]^{-1/2}\left[\hat{X}_\omega(t+\Delta) - X_\omega(t+\Delta)\right] \xrightarrow{d} \mathcal{N}(0,1),$$

*where $\sigma_j^2 := Var(\epsilon_{t_j})$ for all $j$, and*

$$V(t;\Delta,R,\mathbf{\Phi}) := \sum_{j=1}^{J} \sigma_j^2 A_{t_j}^2(t;\Delta,R,\mathbf{\Phi}),$$

$$A_{t_j}(t;\Delta,R,\mathbf{\Phi}) := \sum_{r=0}^{R} \frac{1}{r!}\Delta^r \left\{\mathbf{\Phi}^{(r)}(t)\right\}^\top \left\{\int_0^1 \mathbf{\Phi}(s)\mathbf{\Phi}^\top(s)ds\right\}^{-1} \mathbf{\Phi}(t_j).$$

Theorem 2.2.2 constructs the pointwise asymptotic normality of the functional predictor $\hat{X}_\omega$ through the asymptotically normal prediction error from $\hat{X}_\omega$ to the Bernstein approximating polynomial $B_K(t, X_\omega)$ and the convergence of $B_K(t, X_\omega)$ to $X_\omega$ as $K$ increases towards infinity. Specifically, since the inverse factorial series $\sum_{r=0}^{\infty}(r!)^{-1}$ converges, when $R \to \infty$, we have $\sum_{r=0}^{R}(r!)^{-1}\Delta^r < C < \infty$ for some small $\Delta$ and fixed $C \in \mathbb{R}$, whence

$A_{t_j}(t;\Delta,R,\mathbf{\Phi})$ and thus $A^2_{t_j}(t;\Delta,R,\mathbf{\Phi})$ is bounded. Then by Assumptions 2.2.1 (b) and 2.2.2 (a), the term $V(t;\Delta,R,\mathbf{\Phi})$ is of order $J$. Meanwhile, $\hat{X}_\omega(t+\Delta)-X_\omega(t+\Delta)$ can be decomposed into $\hat{X}_\omega(t+\Delta)-B_K(t+\Delta,X_\omega)$ and $B_K(t+\Delta,X_\omega)-X_\omega(t+\Delta)$, where the former will be inflated to an asymptotic normal, and the latter is dominated because a desirably small approximation error $B_K(t+\Delta,X_\omega)-X_\omega(t+\Delta)$ can be achieved by a sufficiently large $K$. As a result, a $\sqrt{J}$-asymptotic normality can be achieved.

## 2.3 Simulation Analysis

This section presents a simulation study to illustrate the procedure of our FDA-based method. Its performance is then compared to that of a parametric approach under both correct specification and misspecification, thereby revealing FDA's superiority in tracking the f.d. data structures of stochastic processes.

### 2.3.1 The Data-generating Process

The data-generating process is based on the strong GARCH (1,1), where $y_k = (S_k - S_{k-1})/S_{k-1},\ k = 1,2,...$ is the arithmetic return on a financial asset with the price $S_k$. Let $h$ be the time window, and recall that a strong GARCH (1,1) process is represented by $y_k = \mu + \epsilon_k$ with $\epsilon_k \sim \mathcal{N}(0,V_k)$ given the $\sigma$-algebra generated by $\epsilon_{k-1}$ and $V_k = \omega_h + \xi_h\epsilon^2_{k-1} + \gamma_h V_{t_{j-1}}$ for the f.d. parameters $\omega_h,\xi_h,\gamma_h > 0$ and $\xi_h + \gamma_h < 1$ (Engle, 1982; Bollerslev, 1986).

As $h$ shrinks, $(S_{kh},V_{kh})$ weakly converges to its continuous-time limit $(S_t,V_t)$ (Corradi, 2000). This is a simple but motivating example, since given different assumptions on the f.d. parameters' convergence rates (i.e., the convergence of $\xi_h$ can be at rate $\sqrt{h}$ or rate $h$), $(S_t,V_t)$ has been shown to be a diffusion process solution to either a stochastic volatility (SV) model (Nelson, 1990; Meddahi, 2001)

$$\begin{bmatrix} ds_t \\ dv_t \end{bmatrix} = \begin{bmatrix} a \\ \left(\beta - \frac{1}{2}\sigma^2\right) + \alpha\exp(-v_t) \end{bmatrix} dt + \begin{bmatrix} \sqrt{1-\rho^2}\exp(\frac{v_t}{2}) & \rho\exp(\frac{v_t}{2}) \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} dW^1_t \\ dW^2_t \end{bmatrix}, \tag{2.7}$$

where $s_t := \log(S_t)$, $v_t := \log(V_t)$, $W^1_t$ and $W^2_t$ are independent standard Brownian motions, or a deterministic volatility (DV) model (Corradi, 2000)

$$\begin{bmatrix} ds_t \\ dv_t \end{bmatrix} = \begin{bmatrix} a \\ \beta + \alpha\exp(-v_t) \end{bmatrix} dt + \begin{bmatrix} \exp\left(\frac{v_t}{2}\right) \\ 0 \end{bmatrix} dW^1_t. \tag{2.8}$$

Indeed, despite that the ARCH-type diffusion models under both SV and DV have been frequently applied in literature to estimate continuous-time processes (e.g., Das, 2002; Duan et al., 2006; Aït-Sahalia and Kimmel, 2007; Christoffersen et al., 2010; Kleppe et al., 2014; Wu et al., 2018), there has been considerable debate on the choice of the convergence rate assumptions, which makes the parametric analysis such as MLE tailored for either limit questionable (e.g., Wang, 2002; Alexander and Lazar, 2021). More generally, if the f.d. data structures have different limiting processes under different convergence conditions, it would be impractical for parametric continuous-time generalization to exhaust all possible limits to choose the correct one.

Hence we generate pseudo-continuous return and volatility processes as in Equations (2.7) and (2.8) over the interval of $[0, 1]$, consisting of five equally-spaced trading points per day for one year of 252 trading days (i.e., 5 time points per day × 252 days per year × 1 year). The true values of parameters in Table 2.1 are used to generate $N = 1000$ pseudo-continuous log-return and log-volatility[2] trajectories from the SV limit, denoted respectively by $S_S(t)$ and $v_S(t)$, and those trajectories from the DV limit, denoted respectively by $S_D(t)$ and $v_D(t)$. In favor of parametric methods, our simulation study takes the entire pseudo-continuous-time processes as the observations (i.e., $J = 1260$), denoted as $S_S(t_j)$, $v_S(t_j)$, $S_D(t_j)$ and $v_D(t_j)$ for $j = 1, ..., J$, to numerically mimic the scenario where the sampling time window becomes arbitrarily small, and the estimation is performed based on an eight-month rolling window.

### 2.3.2 Fitting and Forecasting with FDA

Now we present the procedures for fitting and forecasting with FDA[3] in steps (a) to (c), and the forecast evaluation is discussed in step (d). Note that, only the implementations in terms of $S_S(t)$ and $S_D(t)$ are explained here, and similar procedures can be applied to $v_S(t)$ and $v_D(t)$ processes.

[2] It should be noted that the volatility is not always observed in practice, and discussions addressing parametric analysis with such unobservability have been provided in literature (e.g., Ledoit and Santa-Clara, 1998; Aït-Sahalia and Kimmel, 2007; Kleppe et al., 2014). Yet, the FDA-based method can handle the return and volatility processes separately; hence in what follows, we assume that volatility is accessible without harming the analysis of the return. Should we be interested in the unobserved volatility, the same FDA-based method can be applied to suitable proxies.

[3] In Appendix B.8, the performance of FDA-based analysis is also illustrated with observations drawn equality spaced on a daily basis, and we show that it is capable of providing reliable forecasts with smaller sample sizes and larger sampling time windows such as a-month-long daily data containing only 21 observations.

(a) For each of the 1000 replications, estimate the underlying processes $S_S(t)$ and $S_D(t)$ as in Equations (2.4) and (2.5) from Section 2.2.1, using order-four B-spline basis[4] where the number of basis functions are selected by a leave-one-out cross validation. We then obtain $\tilde{S}_S(t)$ and $\tilde{S}_D(t)$, respectively, as in Equation (2.5).
(b) Compute the forecasting values $\hat{S}_S(t+\Delta)$ and $\hat{S}_D(t+\Delta)$ using Equation (2.6) with $R = 1$.
(c) Implement the *Kolmogorov-Smirnov* (*K-S*) test on the pairs "$\hat{S}_S(t+\Delta)$ and $S_S(t+\Delta)$", "$\hat{S}_S(t+\Delta)$ and $S_D(t+\Delta)$", "$\hat{S}_D(t+\Delta)$ and $S_S(t+\Delta)$" as well as "$\hat{S}_D(t+\Delta)$ and $S_D(t+\Delta)$" for all time points $t$, to check whether the FDA-based predictors can distinguish the true underlying processes from a falsely assumed continuous-time limit in terms of their distributions.
(d) Calculate the mean-squared forecast error (MSFE) according to Leitch and Tanner (1991) and Hansen and Lunde (2005), between the pairs "$\hat{S}_S(t+\Delta)$ and $S_S(t+\Delta)$", "$\hat{S}_S(t+\Delta)$ and $S_D(t+\Delta)$", "$\hat{S}_D(t+\Delta)$ and $S_S(t+\Delta)$" as well as "$\hat{S}_D(t+\Delta)$ and $S_D(t+\Delta)$" for the out-of-sample performance evaluation.

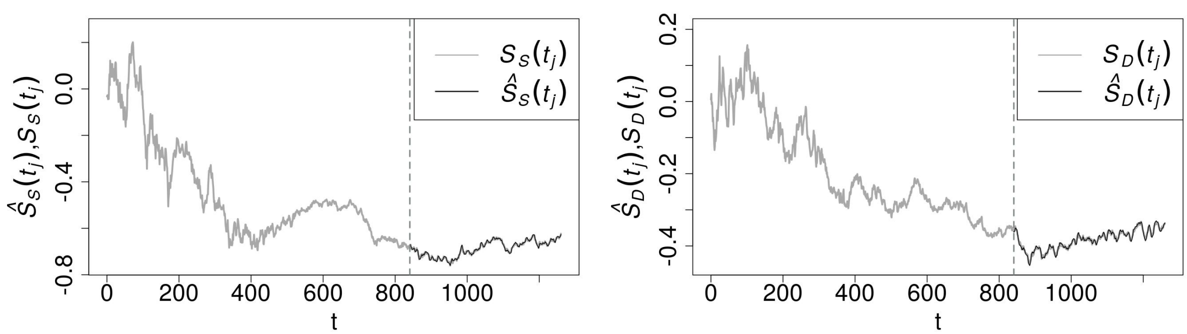


Figure 2.1: Functional Data Prediction, Returns

Figures 2.1 and 2.2 present the prediction results, where the underlying processes are indicated by gray lines and the one-step-ahead rolling forecasts are indicated by the black lines. Then, we use the *K-S* test to compare the predicted and the underlying processes[5].

[4] In practice, a data-driven method can be used to select the optimal order of the B-spline basis and the truncating term $R$ of the Taylor expansion. However, developing such a data-driven method is not the focus of the current paper, so we adopt the commonly used order-four B-spline basis and $R = 1$ (as in step (b)) in our estimation. The tuning parameter and the number of basis functions are selected by a leave-one-out cross validation.

[5] The rejection rate of the null of distribution equality over the entire time domain is shown in Figures B.1 and B.2 of Appendix B.8.

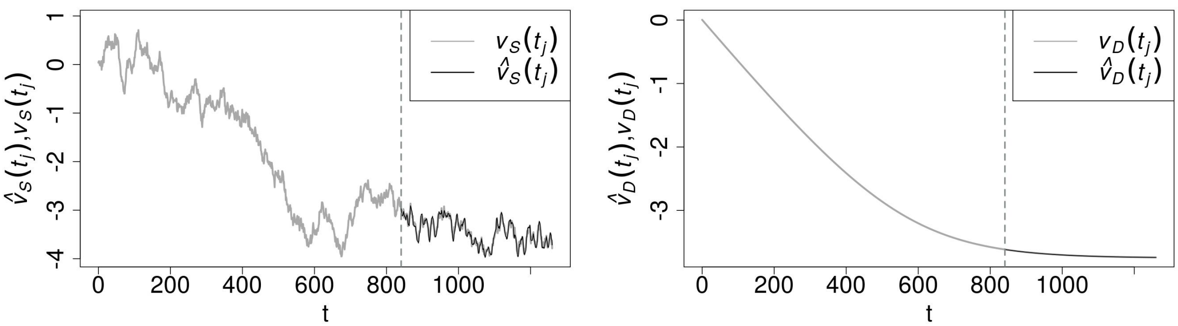


Figure 2.2: Functional Data Prediction, Volatility

We find that, at almost every fixed observing time $t_j$, our forecast data shares the same distribution as its underlying process at the 5% significance level, and the null of distribution equality is rejected (almost) everywhere for the cross-comparisons. The results imply that the one-step-ahead rolling forecast using the FDA-based method can preserve the distributions of the true underlying processes pointwisely and can distinguish the true underlying processes from a falsely assumed continuous-time limit in terms of their distributions.

### 2.3.3 Comparison to Parametric Methods

In this section, we adopt MLE as a representative parametric method for the comparison to our FDA approach because the MLE method has been a commonly used approach for parametrically estimate GARCH-like models (see e.g., Aït-Sahalia 2002; Kleppe et al. 2014). Based on the stochastic differential equations of the two limits in Equations (2.7) and (2.8), the likelihood functions are obtained utilizing the (joint) normality of $dW_t^1$ and $dW_t^2$, and the differentials $ds$ and $dv$ are approximated by the corresponding first difference of the discrete observations.

The results for the first rolling window is shown in Table 2.1 to provide a snapshot of the performance of MLE, while the details for all 420 estimations are plotted in Appendix B.9. With correctly specified models, MLE shows a good asymptotic performance, while the estimates appear to be significantly different from the truth under misspecification. As noted in Table 2.1, the rejection rates for some parameters slightly exceed the reasonable range of errors induced by a binomial distribution with $N = 1000$ trials and a probability of success at 5% even when the model is correctly specified, which can be attributable to the insufficient number of observations. Nevertheless, in general, methods that do not involve such model specifications will be handy to circumvent the discrepancy in continuous-time modeling and analysis.

Table 2.1: Estimating GARCH Parameters with MLE

| Parameters | $a$ | $\alpha$ | $\beta$ | $\sigma$ | $\rho$ |
|---|---|---|---|---|---|
| True values (SV) | 0.1 | 0.2 | -8.5 | 2.7 | -0.8 |
| Fitting SV processes using a SV model (correct specification) | | | | | |
| Estimate | -0.006 | 0.303 | -9.577 | 2.699 | -0.801 |
| Bias | -0.106 | 0.103 | -1.077 | -0.001 | -0.001 |
| Rejection Rate | **0.074** | **0.077** | **0.060** | **0.060** | **0.047** |
| Fitting SV processes using a DV model (misspecification) | | | | | |
| Estimate | 0.101 | 0.476 | -13.563 | – | – |
| Bias | 0.001 | 0.276 | -5.063 | – | – |
| Rejection Rate | **0.063** | **0.808** | **0.906** | – | – |
| True values (DV) | 0.1 | 0.2 | -8.5 | 0 | 0 |
| Fitting DV processes using a DV model (correct specification) | | | | | |
| Estimate | 0.089 | 0.202 | -8.546 | – | – |
| Bias | -0.011 | 0.002 | -0.046 | – | – |
| Rejection Rate | **0.044** | **0.056** | **0.047** | – | – |
| Fitting DV processes using a SV model (misspecification) | | | | | |
| Estimate | 1.068 | 0.200 | -8.500 | 0.000 | 0.163 |
| Bias | 0.968 | 0.000 | -0.000 | 0.000 | 0.163 |
| Rejection Rate | **0.319** | **1.000** | **1.000** | **1.000** | **0.866** |

Note: an 1000-time simulation allows for a "$\pm\sqrt{0.95 * 0.05/1000} * 2 = 0.014$" error on the 5%-rejection rate under correct specification.

Next, the estimation results are used for the one-step-ahead out-of-sample forecasting. Recall Equations (2.7) and (2.8), whence the forecasts based on conditional expectation are such that

$$\begin{cases} \hat{s}_{t+\Delta} := \hat{\mathbb{E}}\left[s_{t+\Delta}|s_t\right] = s_t + \hat{a}\Delta, \\ \hat{v}_{t+\Delta} := \hat{\mathbb{E}}\left[v_{t+\Delta}|v_t\right] = v_t + \left[\left(\hat{\beta} - \frac{1}{2}\hat{\sigma}^2\right) + \hat{\alpha}\exp(-v_t)\right]\Delta \end{cases} \tag{2.9}$$

for the SV limit and

$$\begin{cases} \hat{s}_{t+\Delta} := \hat{\mathbb{E}}\left[s_{t+\Delta}|s_t\right] = s_t + \hat{a}\Delta, \\ \hat{v}_{t+\Delta} := \hat{\mathbb{E}}\left[v_{t+\Delta}|v_t\right] = v_t + \left[\hat{\beta} + \hat{\alpha}\exp(-v_t)\right]\Delta \end{cases} \tag{2.10}$$

for the DV limit, where $\hat{a}, \hat{\beta}, \hat{\alpha}$, and $\hat{\sigma}$ are the estimated parameters from MLE, and $\Delta$ is the forecasting step as previously defined.

Finally, we compare the predictors of the FDA- and MLE-based methods in MSFE. For the prediction of DV volatility using an SV model, since the parameters $\alpha$ and $\beta$ can be estimated accurately without stochastic noise, and the conditional mean prediction relies only on $\alpha$ and $\beta$ as shown in Equation (2.9), this MLE estimation is essentially under correct specification and hence provides a precise forecast. In all three other cases where the MLE is under misspecification, FDA outperforms MLE in terms of MSFE in all 1000 simulations. When MLE is under correct specification, the FDA-based method still dominates MLE in all 1000 simulations for the prediction of the stochastic return and volatility as well as the return of DV process, while 822 out of 1000 times, FDA performs better for the prediction of the DV volatility. The distributions of the comparisons in relative MSFE – the ratio of the MSFEs between FDA and MLE – are presented in Appendix B.8.

## 2.4 Conclusion

In continuous-time modeling, parametric methods fail to provide reliable analysis when there is discrepancy due to the existence of multiple limits. This paper adopts FDA to uncover the true continuous-time underlying processes subject to f.d. data structures under infill asymptotics and suggests a forecasting method by integrating FDA with Taylor series expansion, which also explores an application of FDA in out-of-sample prediction.

Our theorems demonstrate that the FDA-based method only requires the smoothness of the the conditional mean function and the low correlation of the observation errors, and with proper basis expansions, large samples ensure that the functional estimator converges to a unique and well-defined limit. The simulation analysis shows that the FDA method is capable of distinguishing processes with different limits in out-of-sample prediction and outperforms the MLE method in MSFE, which makes it a preferable tool for fitting and forecasting when there is uncertainty in modeling the underlying process. Finally, though we illustrate our idea with a GARCH example, it should be noted that the suggested FDA-based method is not limited to any particular continuous-time model, and can be further extended as in Kearney et al. (2015) to processes with jump dynamics, which are widely used in describing equity returns.

# Chapter 3

# Stochastic Discount Factors in Incomplete Markets[1]

## 3.1 Introduction

Stochastic discount factor (henceforth SDF) forms the basis for all asset pricing and provides a summary of investor preferences for payoffs over different states of the world. Under the law of one price (henceforth LOOP), the asset pricing equation established by Harrison and Kreps (1979), Harrison and Pliska (1981) and Hansen and Jagannathan (1991) implies that asset prices today are a function of their expected future payoffs discounted by the SDF. When markets are complete, the asset pricing equation leads to a unique SDF, whereas there is a multiplicity of admissible SDFs that satisfy the equation in the absence of complete markets (Hansen and Jagannathan, 1991; Boyle et al., 2008; Kaido and White, 2009), thus complicating the study of financial market equilibrium, portfolio optimization, and derivative securities (Skiadas, 2007; Staum, 2007; Boyle et al., 2008). It is therefore essential to establish a framework for characterizing the incomplete market SDF set, and assess the extent of market incompleteness.

Markets are incomplete when perfect risk transfer is impossible, and this incompleteness can be caused by a variety of factors such as jumps or volatility in underlying asset prices (Jackwerth, 2004; Staum, 2007; Willems and Morbee, 2008; Bondarenko and Longarela, 2009; Mnif, 2012; Marroqu et al., 2013; Kwak et al., 2014; Cheridito et al., 2016; Bouzianis and Hughston, 2020). This paper first considers a discrete-time setting, which does not

[1] This chapter is co-authored with Tao Chen

restrict the class of unhedgeable risks, and therefore does not constrain the cause of market incompleteness. Particularly, we regard markets as incomplete when there are extra states relative to the traded assets due to idiosyncratic risks that cannot be diversified by trading the spanning assets in the market. We demonstrate that the incomplete market SDF set constructed under this general setup has a unique boundary point, and only shrinks to this point when the market completes. This nice property allows us to examine features of the incomplete market SDF set, and enables us to determine the degree of market incompleteness.

To facilitate the empirical implementation of our results, we parameterize the market incompleteness in a continuous-time setting and propose that the undiversifiable risk is caused by a specific, but practically realistic source of incompleteness – stochastic jumps, where prices exhibit positive probabilities of unexpected changes, regardless of the interval between successive observations. Jump diffusion processes have been frequently used to model asset pricing, and their empirical performance in fitting the time-series properties of the asset price has been extensively evidenced by a number of studies (Dritschel and Protter, 1999; Svishchuk et al., 2000; Bellamy, 2001; Andersen et al., 2002; Carr et al., 2002; Geman, 2002; Willems and Morbee, 2008; Bouzianis and Hughston, 2020; Aït-Sahalia et al., 2021). In most cases, jumps cause incompleteness, except in very simple or unusual models, whence the market offers sufficient trading opportunities (Dritschel and Protter, 1999; Staum, 2007). As such, inspired by Merton (1976)'s work, whereby the total change in price should be a combination of the normal and abnormal price vibrations, our continuous-time framework considers complete markets as those in which asset prices are subject only to normal fluctuations, and incomplete markets as those with a positive likelihood of experiencing unanticipated changes in price. We demonstrate that the theoretical results developed in the discrete-time counterpart are still valid in the continuous-time setting, and we further use those results to establish the degree of market incompleteness.

In the literature, one popular measure for the degree of market incompleteness is through the correlation between the derivative price and its basis asset values (Cass and Citanna, 1998; Marin and Rahi, 2000; Dávila et al., 2017; Chen et al., 2021), where a lower correlation indicates a greater degree of incompleteness, and the market is complete only when the correlation reaches 100%. Another measure employs the root-mean-squared error between the payoff function of the derivative and the value of the optimal-replication portfolio constructed by the underlying securities (Bertsimas et al., 2001). The degree of incompleteness is thus determined by the extent to which the replicated portfolio is able to correctly price the derivative of the underlying assets.

Our approach is distinct from the previous ones in that instead of focusing on the linkage between the prices of derivative securities and their underlying assets, we only concern the prices of the primitive assets. In particular, considering that SDFs summarize investor preferences for payoffs across different states of the world, we define the degree of market incompleteness as how much the investor's risk preference under incomplete markets diverges from that under complete markets. The empirical implementation of this measure is summarized as follows, and we will elaborate in Sections 3.2 and 3.3 with discrete- and continuous-time examples. After constructing the incomplete market SDF set and determining its corresponding complete SDF boundary point using the asset prices, we employ the distance between their probability distributions to estimate the degree of market incompleteness. As the complete market SDF is the boundary point of the incomplete market SDF set, this distance vanishes only if the extra-state probability is zero, that is, when the incomplete market SDF set degenerates into a unique complete market SDF. It can be challenging to gauge this distance, since the complete and incomplete SDFs have probability distributions of different dimensions, i.e., there are extra states with positive probabilities under incomplete markets compared with complete markets. A natural solution to this problem is the Wasserstein metric, a widely adopted measure in estimating the distance between distributions whose supports differ, and its value reflects the least cost required to transform from one distribution to another (Mallows, 1972; Del Barrio et al., 1999; Villani, 2009; Nguyen, 2011).

The remainder of this paper proceeds as follows. Sections 3.2 and 3.3 sketch the discrete- and continuous-time frameworks to model SDF under incomplete markets, and show the applicability of our model in assessing the evolution of market incompleteness. Section 3.4 provides the empirical analysis and investigates the evolution of incompleteness in four of the world's largest stock markets. Section 3.5 concludes with a summary and a discussion of directions for future research.

## 3.2 Discrete-time Setting

In this section, we model the SDF set and the market incompleteness under three discrete-time setups, where each case is denoted as one risk-free bond$-A$ risky asset(s)$-T$ periods$-S$ states with $A \in \mathbb{N}$, and $T \geq 2 \in \mathbb{N}$. The number of traded assets is assumed to be less than the number of states at the end of each period, i.e., $A + 1 < S$, so that the markets are incomplete, while the cause of this incompleteness is not imposed. Then, there is a set of SDFs identified by the distribution of observed asset prices (Boyle et al., 2008; Kaido and White, 2009).

To motivate our study, we begin with a two-period market with only one additional state relative to the number of traded assets in Sections 3.2.1 and 3.2.2. We formalize the setup as follows.

**Assumption 3.2.1.** *Suppose that there are one risk-free bond and $A \in \mathbb{N}$ risky assets. We consider a two-period market, $t \in \{0,1\}$, with trading occurring on dates $t = 0,1$. The outcome of the second period, $t = 1$, is uncertain, and is represented by a finite set $\Omega = \{\omega^s\}_{s=1,2,\ldots,S}$ comprising $S = A+2$ states of nature. Let $\mathcal{F}$ be the set of events with all subsets of $\Omega$ and $P$ be the physical probability measure such that $P : \mathcal{F} \to \mathbb{R}$. There exists a set $\mathcal{P}$ of complete probability measures on $(\Omega, \mathcal{F})$ such that $P \in \mathcal{P}$. Letting $P(\omega^s) = \pi^s$ be the probability of state $\omega^s$ such that $\pi^s$ are strictly positive scalars for all $s = 1,2,\ldots,S$ in incomplete markets, while $\pi^S = 0$ in complete markets.*

Assumption 3.2.1 has three implications. First, there are only two periods in the economy, and thus, we do not index the states by time in the subsequent two sections. We will extend our setup to multiperiods, where $t$ takes the value from a finite sequence of real numbers in $[0,1]$ that are equally-spaced, and continuous-time, where $t$ is generalized to take any value in $[0,1]$. Second, without loss of generality, the last state is assumed to be the extra one, which is caused by an unknown source of idiosyncratic risk that cannot be hedged by the existing marketed assets. Third, our basic design requires the markets to be either complete, with the same number of marketed assets and states, or incomplete, with only one extra state. In Section 3.2.4, this restriction is relaxed to a finite number of extra states, and $\pi^S$ is extended to a vector such that[2] $[\boldsymbol{\pi}^{\bar{s}}]_{\bar{s}=A+2,A+3,\ldots,S} \in \mathbb{R}^{S-1-A}_{++}$. Then, the market completes only when[3] $[\boldsymbol{\pi}^{\bar{s}}]_{\bar{s}=A+2,A+3,\ldots,S} = \mathbf{0}_{S-1-A}$.

### 3.2.1 One Risk-free Bond, One Risky Asset, Two periods, Three States (1-1-2-3)

Suppose that we have one risk-free bond and one risky primitive asset in the economy, and there are three states at period $t = 1$ such that $\Omega = (\omega^1, \omega^2, \omega^3)$, correspondingly there exists a set of physical probabilities

$$\boldsymbol{\Pi} = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \pi^3\right]^\top \in \mathbb{R}^3_{++} : \sum_{s=1}^{3} \pi^s = 1 \right\}. \tag{3.1}$$

As we have two assets, the gross rate of returns realized at the second period are of length two[4], i.e., $\boldsymbol{r}(\omega^s) = [r^s, 1]^\top$, where $r^s$ denotes the return for the risky asset in state

[2] We write $\boldsymbol{v} \in \mathbb{R}^n_{++}$ for a vector that is strictly positive in all its coordinates.

[3] $\mathbf{0}_n$ denotes a zero vector of size $n$.

[4] For vectors and matrices, we shall use the superscript '$\top$' to denote transpose.

$s$ and, for simplicity, the return of the bond is 1, suggesting a zero risk-free rate. Let $\boldsymbol{r} = [\boldsymbol{r}(\omega^1), \boldsymbol{r}(\omega^2), \boldsymbol{r}(\omega^3)]$, we assume that the second-moment matrix of $\boldsymbol{r}$, $\mathrm{E}[\boldsymbol{r}\boldsymbol{r}^\top]$, is nonsingular, so that the cases where the entries of $\boldsymbol{r}$ is linearly dependent are ruled out. This restriction also guarantees that LOOP holds trivially for linear combinations of $\boldsymbol{r}$ (Hansen and Jagannathan, 1991). We can treat $\boldsymbol{r}$ as payoffs for assets with price one, and the asset pricing equation is expressed as[5]

$$\mathrm{E}_{\boldsymbol{\pi}}[\boldsymbol{r}\boldsymbol{M}] = \sum_{s=1}^{3} \boldsymbol{r}(\omega^s)\, \boldsymbol{M}(\omega^s)\, \pi^s = \mathbf{1}_2, \tag{3.2}$$

where the subscript of E is used to specify which probability measure is being used to compute the expectation. As discussed in Kaido and White (2009), the SDF $\boldsymbol{M}$ is a non-zero $\mathcal{F}$-measurable random variable such that $\boldsymbol{M} : \Omega \to \mathbb{M}_{\boldsymbol{\pi}}$, where $\mathbb{M}_{\boldsymbol{\pi}}$ is the set of SDFs under $P \in \mathcal{P}$ that satisfies Equation (3.2):

$$\mathbb{M}_{\boldsymbol{\pi}} := \{\boldsymbol{M} : \mathrm{E}_{\boldsymbol{\pi}}[\boldsymbol{r}\boldsymbol{M}] = \mathbf{1}_2\}. \tag{3.3}$$

Let $M^s \equiv \boldsymbol{M}(\omega^s)$ and $\alpha = \pi^3 M^3 \in \mathbb{R}_*$ be the free variable[6,7], for any $\boldsymbol{\pi} \in \boldsymbol{\Pi}$, we can think of $\boldsymbol{M}$ as a vector in $\mathbb{R}^3_*$, where the three coordinates give the values of $\boldsymbol{M}$ on the three possible outcomes. Thus, Equation (3.3) can be rewritten as[8]

$$\mathbb{M}_{\boldsymbol{\pi}} = \left\{ \boldsymbol{M} \in \mathbb{R}^3_* : \begin{bmatrix} M^1 \\ M^2 \\ M^3 \end{bmatrix} = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \\ 0 \end{bmatrix} \boldsymbol{\pi}^{-1} + \alpha \begin{bmatrix} \frac{r^2-r^3}{r^1-r^2} \\ -\frac{r^1-r^3}{r^1-r^2} \\ 1 \end{bmatrix} \boldsymbol{\pi}^{-1}, \alpha \in \mathbb{R}_* \right\}.$$

Lastly, we write $\boldsymbol{C}$ as the set that combines all $\mathbb{M}_{\boldsymbol{\pi}}$'s given $\boldsymbol{\pi} \in \boldsymbol{\Pi}$ such that $\boldsymbol{C} := \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \boldsymbol{\Pi}\}$.

[5] $\mathbf{1}_n$ is a vector of ones in $\mathbb{R}^n$.

[6] $\boldsymbol{M}$ can be also thought as the discounted Radon–Nikodym derivative, where the Radon–Nikodym derivative $\mathcal{D}$ is defined as a $\mathcal{F}$-measurable random variable such that for any $A \in \mathcal{F}, Q(A) = \int_A \mathcal{D} dP$ (Kaido and White, 2009) with $Q$ being the risk-neutral probability measure. In our setup, assuming zero risk-free interest rate, $\boldsymbol{M} = \mathrm{E}[dQ/dP|\mathcal{F}]$. Since $Q$ and $P$ are equivalent in measure, they agree on which events have zero probability, and hence, $\boldsymbol{M}$ is non-zero.

[7] We write $\boldsymbol{v} \in \mathbb{R}^n_*$ for a vector that is non-zero in all its coordinates.

[8] To simplify our notation, for two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same sizes, $\boldsymbol{AB}$ is their element-wise product with the same size, and its element is expressed as $(\boldsymbol{AB})_i = \boldsymbol{A}_i \times \boldsymbol{B}_i$. Similarly, for a vector $\boldsymbol{A}$, the element-wise power of a real number $x$ on it is $\boldsymbol{A}^x$, i.e., $(\boldsymbol{A}^x)_i = (\boldsymbol{A}_i)^x$; for a real number $x$, the element-wise power of a vector $\boldsymbol{A}$ on it is $x^{\boldsymbol{A}}$, i.e., $(x^{\boldsymbol{A}})_i = x^{\boldsymbol{A}_i}$.

### Set properties of $\boldsymbol{C}$

The following proposition provides the limit and boundary points of the set of probability measures in Equation (3.1), which will later be used to explore the boundary point of the incomplete market SDF set $\boldsymbol{C}$. The proof is in Appendix C.2.

**Proposition 3.2.1.** *Consider the metric space* $(\bar{\boldsymbol{\Pi}}, d)$ *such that*

$$\bar{\boldsymbol{\Pi}} = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \pi^3\right]^\top : \sum_{s=1}^{3} \pi^s = 1, \pi^1, \pi^2 > 0, \pi^3 \geq 0 \right\}$$

*and* $d$ *is the Euclidean distance metric. Then, the set of limit points of* $\boldsymbol{\Pi}$ *in* $(\bar{\boldsymbol{\Pi}}, d)$ *is*

$$L(\boldsymbol{\Pi}) = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \pi^3\right]^\top : \sum_{s=1}^{3} \pi^s = 1, \pi^1, \pi^2 > 0, \pi^3 \geq 0 \right\}$$

*and the set of boundary points of* $\boldsymbol{\Pi}$ *in* $\bar{\boldsymbol{\Pi}}$ *is*

$$\partial \boldsymbol{\Pi} = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \pi^3\right]^\top : \sum_{s=1}^{3} \pi^s = 1, \pi^1, \pi^2 > 0, \pi^3 = 0 \right\}.$$

The probability set $\bar{\boldsymbol{\Pi}}$ has the first two states being strictly positive and the last state being nonnegative, which therefore, covers all complete and incomplete market scenarios described in Assumption 3.2.1. Proposition 3.2.1 implies that there is a unique boundary point for $\boldsymbol{\Pi}$ in $(\bar{\boldsymbol{\Pi}}, d)$ when $\lim \pi^3 \to 0$, which is compatible with Assumption 3.2.1 such that the incompleteness is introduced through a non-tradable risk with positive likelihood of occurrence.

In accordance to Proposition 3.2.1, the next result presents that the combined incomplete market SDF set $\boldsymbol{C}$ has the complete market SDF on its boundary, and its proof can be found in Appendix C.3.

**Theorem 3.2.1.** *Consider a metric space* $\left(\bar{\boldsymbol{C}}, d_1\right)$ *with* $\bar{\boldsymbol{C}} \coloneqq \left\{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \bar{\boldsymbol{\Pi}}\right\}$ *and* $d_1$ *being the Wasserstein distance such that for* $\boldsymbol{x}, \boldsymbol{y} \in \bar{\boldsymbol{\Pi}}$,

$$d_1\left(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}\right) = \inf_{\boldsymbol{w}} \left\{ \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2\left(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}\right) : \boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y}) \right\}, \tag{3.4}$$

*where*[9] $W(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \left\{\boldsymbol{w} \in \mathbb{R}_+^{S_{\boldsymbol{y}} \times S_{\boldsymbol{x}}} : \boldsymbol{w}^\top \mathbf{1}_{S_{\boldsymbol{y}}} = \boldsymbol{x}, \boldsymbol{w}\boldsymbol{x} = \boldsymbol{y}\right\}$ *is the set of transport plans between* $\boldsymbol{x}$ *and* $\boldsymbol{y}$. $S_{\boldsymbol{z}}$ *is the number of states with non-zero probabilities and the superscript* $s_{\boldsymbol{z}}$

[9] We write $A \in \mathbb{R}_+^{M \times N}$ for matrix of dimension $M \times N$ that is non-negative in all its elements.

*is the index of the elements in the vector under the physical probability* $\boldsymbol{z} \in \bar{\boldsymbol{\Pi}}$. *For all* $s_{\boldsymbol{x}} = 1, 2, \ldots, S_{\boldsymbol{x}}$ *and* $s_{\boldsymbol{y}} = 1, 2, \ldots, S_{\boldsymbol{y}}$,

$$d_2\left(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}\right) = \left|v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}}\right| + \left|u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}}\right|, \tag{3.5}$$

*where*

$$\begin{cases} \boldsymbol{v}(\boldsymbol{x}) = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \\ 0 \end{bmatrix} \boldsymbol{x}^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}) = \begin{bmatrix} \frac{r^2-r^3}{r^1-r^2} \\ -\frac{r^1-r^3}{r^1-r^2} \\ 1 \end{bmatrix} \boldsymbol{x}^{-1}, & \boldsymbol{x} \in \boldsymbol{\Pi}; \\ \boldsymbol{v}(\boldsymbol{x}) = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \end{bmatrix} \boldsymbol{x}^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \boldsymbol{x}^{-1}, & \boldsymbol{x} \in \partial\boldsymbol{\Pi}. \end{cases}$$

*Then, the set of limit points of* $\boldsymbol{C}$ *in* $(\bar{\boldsymbol{C}}, d_1)$ *is*

$$L(\boldsymbol{C}) = \left\{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \boldsymbol{\Pi}\right\} \cup \left\{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \partial\boldsymbol{\Pi}\right\},$$

*where, for any* $\boldsymbol{\pi} \in \partial\boldsymbol{\Pi}$,

$$\mathbb{M}_{\boldsymbol{\pi}} = \left\{ \begin{bmatrix} M^1 \\ M^2 \end{bmatrix} = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \end{bmatrix} \boldsymbol{\pi}^{-1} \right\},$$

*and the set of boundary points of* $\boldsymbol{C}$ *is denoted as* $\partial\boldsymbol{C} = \left\{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \partial\boldsymbol{\Pi}\right\}$.

Theorem 3.2.1 utilizes the Wasserstein metric as the distance measure, which is a natural way to compare two probability distributions with different supports, and thus, suitable to quantify the divergence of the incomplete market SDFs from the complete market one. Then, we have the following Lemma 3.2.1, suggesting that, for every incomplete market SDF set, $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, there is a complete market SDF $\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}$ that minimizes the distance between them. The proof of this lemma is presented in Appendix C.4, and we will further utilize it in the discussion of set properties of $\boldsymbol{C}$ and the measure for market incompleteness.

**Lemma 3.2.1.** *For every* $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, *there exists* $\mathbb{M}_{\boldsymbol{y}^*}$ *such that*

$$\mathbb{M}_{\boldsymbol{y}^*} = \underset{\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}}{\arg\min}\ d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}).$$

The next result develops an overview of the incomplete market SDF set, where $\boldsymbol{C}$ is *convex*, *open*, *bounded*, and *not compact*. The proof is provided in Appendix C.5.

**Theorem 3.2.2.** *Let $\boldsymbol{\Pi}$ be the set of physical probabilities satisfying Equation* (3.1).

*Let $\mathbb{M}_{\boldsymbol{\pi}}$ be the identified SDF set in the 1-1-2-3 case satisfying Equation* (3.3) *under $\boldsymbol{\pi} \in \boldsymbol{\Pi}$.*

*Let $\boldsymbol{C}$ be the combined SDF set such that $\boldsymbol{C} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \boldsymbol{\Pi}\}$. Then, $\boldsymbol{C}$ is a convex set.*

*Let $\left(\bar{\boldsymbol{C}}, d_1\right)$ be the metric space such that $\bar{\boldsymbol{C}} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \bar{\boldsymbol{\Pi}}\}$ and for $\boldsymbol{x}, \boldsymbol{y} \in \bar{\boldsymbol{\Pi}}$, $d_1$ is as defined in Equation* (3.4). *Then, $\boldsymbol{C}$ is open, bounded and not compact under $\left(\bar{\boldsymbol{C}}, d_1\right)$.*

#### Measure for market incompleteness

As discussed in Theorems 3.2.1 and 3.2.2, we can hence employ the metric $d_1$ defined in Equation (3.4) to measure for market incompleteness. Assuming that, at $t = 0$, the complete and incomplete market SDFs are 1, and the distance between them is 0 following the metric $d_1$, for every $\boldsymbol{x} \in \bar{\boldsymbol{\Pi}}$, the degree of market incompleteness measured at $t = 1$ is defined as in Equation (3.6), which is the least transport cost from the SDF set $\mathbb{M}_{\boldsymbol{x}} \in \bar{\boldsymbol{C}}$ to the complete market SDF set $\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}$ :

$$\begin{aligned} MI(\boldsymbol{x}) &= \min_{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) \\ &= \min_{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}} \inf_{\boldsymbol{w}} \left\{ \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2\left(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}\right) : \boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y}) \right\}. \end{aligned} \tag{3.6}$$

Let $\mathbb{M}^*_{\boldsymbol{y}} = \arg\min_{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$, since $d_1$ is a valid metric and the complete market SDF is the boundary point of the incomplete market SDF set, the degree of market incompleteness equals zero only when markets are complete, i.e., $\lim_{\mathbb{M}_{\boldsymbol{x}} \to \mathbb{M}_{\boldsymbol{y}^*}} MI(\boldsymbol{x}) = 0$. As this degree increases (decreases), the cost to transport the incomplete market SDF to the complete market SDF increases (decreases), which implies more (less) divergence of the current market from completeness.

### 3.2.2 One Risk-free Bond, $A$ Risky Assets, Two Periods, $A + 2$ States (1-$A$-2-$A + 2$)

This section extends the previous economy by having $A \geq 2$ risky primitive assets and $A+2$ states at $t = 1$ such that $\Omega = (\omega^s)_{s=1,2,\ldots,A+2}$. Correspondingly, for $P \in \mathcal{P}$, there is a set of physical probabilities

$$\boldsymbol{\Pi} = \left\{ \left[\pi^1, \pi^2, \ldots, \pi^{A+2}\right]^\top \in \mathbb{R}^{A+2}_{++} : \sum_{s=1}^{A+2} \pi^s = 1 \right\}. \tag{3.7}$$

In each state $s$, assuming a zero risk-free interest rate, the gross rate of return vector realized is of length $A+1$ and denoted as $\boldsymbol{r}(\omega^s) = \left[r^{1,s}, r^{2,s}, \ldots, r^{A,s}, 1\right]^\top$, where $r^{a,s}$ denotes the gross rate of return for the risky asset $a$, and we still assume a zero risk-free rate. Let $\boldsymbol{r} = [\boldsymbol{r}(\omega^1), \boldsymbol{r}(\omega^2), \ldots, \boldsymbol{r}(\omega^{A+2})]$, the second-moment matrix of $\boldsymbol{r}$ is again nonsingular. Recall that based on the asset pricing equation

$$\mathrm{E}_{\boldsymbol{\pi}}[\boldsymbol{r}\boldsymbol{M}] = \sum_{s=1}^{A+2} \boldsymbol{r}\left(\omega^s\right) \boldsymbol{M}\left(\omega^s\right) \pi^s = \mathbf{1}_{A+1}, \tag{3.8}$$

the SDF is a non-zero $\mathcal{F}$-measurable random variable such that $\boldsymbol{M} : \Omega \to \mathbb{M}_{\boldsymbol{\pi}}$, where $\mathbb{M}_{\boldsymbol{\pi}}$ is the set of SDFs under $P \in \mathcal{P}$ that satisfies Equation (3.8):

$$\mathbb{M}_{\boldsymbol{\pi}} = \left\{\boldsymbol{M} : \mathrm{E}_{\boldsymbol{\pi}}[\boldsymbol{r}\boldsymbol{M}] = \mathbf{1}_{A+1}\right\}. \tag{3.9}$$

Let $M^s \equiv \boldsymbol{M}(\omega^s)$ and $\alpha = \pi^{A+2} M^{A+2} \in \mathbb{R}_*$ be the free variable, for any $\boldsymbol{\pi} \in \boldsymbol{\Pi}$, we can think of $\boldsymbol{M}$ as a vector in $\mathbb{R}_*^{A+2}$, where each coordinate gives the value of $\boldsymbol{M}$ on the corresponding outcome. Thus, Equation (3.9) implies that[10]

$$\boldsymbol{r}' \boldsymbol{M}_{(1:A+1)*} \boldsymbol{\pi}_{(1:A+1)*} = \mathbf{1}_{A+1} + \alpha\left(-\boldsymbol{r}''\right),$$

where $\boldsymbol{r}' = (\boldsymbol{r})_{*(1:A+1)}$ and $\boldsymbol{r}'' = (\boldsymbol{r})_{*(A+2)}$, and we can be rewrite it as

$$\mathbb{M}_{\boldsymbol{\pi}} = \left\{\boldsymbol{M} \in \mathbb{R}_*^{A+2} : \boldsymbol{M} = \boldsymbol{v}(\boldsymbol{\pi}) + \alpha \boldsymbol{u}(\boldsymbol{\pi}), \alpha \in \mathbb{R}_*\right\},$$

where

$$\boldsymbol{v}(\boldsymbol{\pi}) = \begin{bmatrix} \left(\boldsymbol{r}'\right)^{-1} \mathbf{1}_{A+1} \\ 0 \end{bmatrix} \boldsymbol{\pi}^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{\pi}) = \begin{bmatrix} -\left(\boldsymbol{r}'\right)^{-1}\left(\boldsymbol{r}''\right) \\ 1 \end{bmatrix} \boldsymbol{\pi}^{-1}.$$

Lastly, the combined set $\boldsymbol{C}$ of $\mathbb{M}_{\boldsymbol{\pi}}$'s for all $\boldsymbol{\pi} \in \boldsymbol{\Pi}$ is defined as $\boldsymbol{C} := \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \boldsymbol{\Pi}\}$.

### Set properties of $\boldsymbol{C}$

As in the 1-1-2-3 case, we start by showing that the complete market SDF is indeed the boundary point of the incomplete market SDF set. The following proposition demonstrates that the probability distribution under complete markets is the boundary point of the set of probabilities under incomplete markets, and its proof is discussed in Appendix C.6.

[10] For a matrix $\boldsymbol{A}$, the $i^{th}$ row of the matrix is denoted as $\boldsymbol{A}_{i*}$ and the $j^{th}$ column of the matrix is denoted as $\boldsymbol{A}_{*j}$.

**Proposition 3.2.2.** *Consider the metric space* $(\bar{\mathbf{\Pi}}, d)$ *such that*

$$\bar{\mathbf{\Pi}} = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \ldots, \pi^{A+2}\right]^\top : \sum_{s=1}^{A+2} \pi^s = 1, \pi^s > 0 \text{ for } s = 1, 2, \ldots, A+1, \pi^{A+2} \geq 0 \right\}$$

*and* $d$ *is the Euclidean distance metric. Then, the set of limit points of* $\mathbf{\Pi}$ *in* $(\bar{\mathbf{\Pi}}, d)$ *is*

$$L(\mathbf{\Pi}) = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \ldots, \pi^{A+2}\right]^\top : \sum_{s=1}^{3} \pi^s = 1, \pi^s > 0 \text{ for } s = 1, 2, \ldots, A+1, \pi^{A+2} \geq 0 \right\}$$

*and the set of boundary points of* $\mathbf{\Pi}$ *in* $\bar{\mathbf{\Pi}}$ *is*

$$\partial\mathbf{\Pi} = \left\{ \boldsymbol{\pi} = \left[\pi^1, \pi^2, \ldots, \pi^{A+2}\right]^\top : \sum_{s=1}^{A+2} \pi^s = 1, \pi^s > 0 \text{ for } s = 1, 2, \ldots, A+1, \pi^{A+2} = 0 \right\}.$$

Then, the following result corroborates with Theorem 3.2.1 that the constructed incomplete SDF set indeed has its boundary point to be the complete market SDF under the defined metric space. The proof of this result is presented in Appendix C.7

**Theorem 3.2.3.** *Consider the metric space* $\left(\bar{\boldsymbol{C}}, d_1\right)$ *with* $\bar{\boldsymbol{C}} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \bar{\mathbf{\Pi}}\}$ *and* $d_1$ *being the Wasserstein distance such that for* $\boldsymbol{x}, \boldsymbol{y} \in \bar{\mathbf{\Pi}}$,

$$d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = \inf_{\boldsymbol{w}} \left\{ \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2\left(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}\right) : \boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y}) \right\}, \tag{3.10}$$

*where* $W(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \{\boldsymbol{w} \in \mathbb{R}_+^{S_{\boldsymbol{y}} \times S_{\boldsymbol{x}}} : \boldsymbol{w}^\top \mathbf{1}_{S_{\boldsymbol{y}}} = \boldsymbol{x}, \boldsymbol{w}\boldsymbol{x} = \boldsymbol{y}\}$ *is the set of transport plans between* $\boldsymbol{x}$ *and* $\boldsymbol{y}$. $S_{\boldsymbol{z}}$ *is the number of states with non-zero probabilities and the subscript* $s_{\boldsymbol{z}}$ *is the index of the elements in the vector under the physical probability* $\boldsymbol{z} \in \bar{\mathbf{\Pi}}$. *For all* $s_{\boldsymbol{x}} = 1, 2, \ldots, S_{\boldsymbol{x}}$ *and* $s_{\boldsymbol{y}} = 1, 2, \ldots, S_{\boldsymbol{y}}$,

$$d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = \left|v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}}\right| + \left|u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}}\right|,$$

*where*

$$\begin{cases} \boldsymbol{v}(\boldsymbol{x}) = \begin{bmatrix} (\boldsymbol{r}')^{-1} \mathbf{1}_{A+1} \\ 0 \end{bmatrix} \boldsymbol{x}^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}) = \begin{bmatrix} -(\boldsymbol{r}')^{-1} (\boldsymbol{r}'') \\ 1 \end{bmatrix} \boldsymbol{x}^{-1}, & \boldsymbol{x} \in \mathbf{\Pi}; \\ \boldsymbol{v}(\boldsymbol{x}) = (\boldsymbol{r}')^{-1} \mathbf{1}_{A+1} \boldsymbol{x}^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}) = \mathbf{0}_{A+1} \boldsymbol{x}^{-1}, & \boldsymbol{x} \in \partial\mathbf{\Pi}. \end{cases}$$

*Then, the set of limit points of* $\boldsymbol{C}$ *in* $(\bar{\boldsymbol{C}}, d_1)$ *can be denoted as*

$$L(\boldsymbol{C}) = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \mathbf{\Pi}\} \cup \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \partial\mathbf{\Pi}\},$$

*where, for any* $\boldsymbol{\pi} \in \partial\boldsymbol{\Pi}$,

$$\mathbb{M}_{\boldsymbol{\pi}} = \left\{ \boldsymbol{M} = \left[ \left(\boldsymbol{r}'\right)^{-1} \mathbf{1}_{A+1} \boldsymbol{\pi}^{-1}_{1:(A+1)} \right] \right\},$$

*and the set of boundary points of* $\boldsymbol{C}$ *is then* $\partial\boldsymbol{C} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \partial\boldsymbol{\Pi}\}$.

Based on Theorem 3.2.3, we can then derive the following lemma, which will be employed further in the discussion of set properties and the degree of market incompleteness. Its proof is shown in Appendix C.8

**Lemma 3.2.2.** *For every* $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, *there exists* $\mathbb{M}_{\boldsymbol{y}^*}$ *such that*

$$\mathbb{M}_{\boldsymbol{y}^*} = \underset{\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}}{\arg\min}\ d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}).$$

The next result establishes the *convexity*, *openness*, *boundedness*, and *non-compactness*, for $\boldsymbol{C}$ in the 1-$A$-2-($A$+2) case with its proof discussed in Appendix C.9

**Theorem 3.2.4.** *Let* $\boldsymbol{\Pi}$ *be the set of all the probability density measures under* $P$ *satisfying Equation* (3.7).

*Let* $\mathbb{M}_{\boldsymbol{\pi}}$ *be the identified SDF set in the 1-A-2-(A+2) case satisfying Equation* (3.9) *given* $\boldsymbol{\pi}$ *in* $\boldsymbol{\Pi}$.

*Let* $\boldsymbol{C}$ *be the combined SDF set such that* $\boldsymbol{C} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \boldsymbol{\Pi}\}$. *Then,* $\boldsymbol{C}$ *is a convex set.*

*Let* $\left(\bar{\boldsymbol{C}}, d_1\right)$ *be the metric space such that* $\bar{\boldsymbol{C}} = \{\mathbb{M}_{\boldsymbol{\pi}}, \boldsymbol{\pi} \in \bar{\boldsymbol{\Pi}}\}$ *and for* $\boldsymbol{x}, \boldsymbol{y} \in \bar{\boldsymbol{\Pi}}$, $d_1$ *is as defined in Equation* (3.10). *Then,* $\boldsymbol{C}$ *is open, bounded and not compact under* $\left(\bar{\boldsymbol{C}}, d_1\right)$.

### Measure for market incompleteness

Similar to the 1-1-2-3 case, based upon Theorems 3.2.3 and 3.2.4, we adopt $d_1$ defined in Equation (3.10) as the measure for market incompleteness. Given that at $t = 0$, the complete and incomplete market SDFs are assumed to be 1, and the distance between them is 0 following the metric $d_1$, for every $\boldsymbol{x} \in \bar{\boldsymbol{\Pi}}$, the degree of market incompleteness is defined as in Equation (3.11), which is the least transport cost from $\mathbb{M}_{\boldsymbol{x}} \in \bar{\boldsymbol{C}}$ to $\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}$:

$$MI(\boldsymbol{x}) = \min_{\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}). \tag{3.11}$$

Let $\mathbb{M}^*_{\boldsymbol{y}} = \arg\min_{\mathbb{M}_{\boldsymbol{y}} \in \partial\boldsymbol{C}} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$, since $d_1$ is a valid metric and the complete market SDF is the boundary point of the incomplete market SDF set, the degree of market incompleteness equals zero only when the markets become complete, i.e., $\lim_{\mathbb{M}_{\boldsymbol{x}} \to \mathbb{M}_{\boldsymbol{y}^*}} MI(\boldsymbol{x}) = 0$. A higher (lower) degree suggests more (less) transport cost is required from the incomplete to the complete market SDF, implying that the market is further away from (closer to) the complete market.

### 3.2.3 One Risk-free Bond, One Risky Asset, Three Periods, Three States (1-1-3-3)

We now extend our layout to a three-period financial market. Consider a time interval $[0,1]$, there are 2 equally-spaced subperiods in $[0,1]$, and $h = 1/2$ is the time window. Suppose we have two long-lived assets, one risk-free bond and one risky primitive asset, available for trading at time points $\{0, 1/2, 1\}$, and three states at each $t = \{kh\}_{k=1,2}$ such that the finite set of states $\Omega_t = (\omega_t^1, \omega_t^2, \omega_t^3)$. Letting $P_t(\omega_t^s) = \pi_t^s$ be the physical probability of state $\omega_t^s$, the corresponding set of physical probabilities under $P_t$ is

$$\boldsymbol{\Pi}_t = \left\{ \boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \pi_t^3\right]^\top \in \mathbb{R}^3_{++} : \sum_{s=1}^{3} \pi_t^s = 1 \right\}. \tag{3.12}$$

Assuming a zero risk-free interest rate, the gross rate of asset returns realized at $t = kh$ in state $s$ is of length two and denoted as $\boldsymbol{r}_t(\omega_t^s) = [r_t^{1,s}, 1]^\top$, where $r_t^{a,s}$ is the return of the risky asset $a$ in state $s$ at time $t$. Let $\boldsymbol{r}_t = [\boldsymbol{r}_t(\omega_t^1), \boldsymbol{r}_t(\omega_t^2), \boldsymbol{r}_t(\omega_t^3)]$, we assume as in the previous sections that the second-moment matrix of $\boldsymbol{r}_t$ is nonsingular. Then, each subperiod $[(k-1)h, kh]$ can be viewed as a two-period model as in the 1-1-2-3 case, and we have the random variable $\boldsymbol{m}_t : \Omega_t \to \mathbb{m}_{\boldsymbol{\pi}_t}$, where $\mathbb{m}_{\boldsymbol{\pi}_t}$ is the set of subperiod SDFs under $P_t \in \mathcal{P}_t$ that satisfies the asset pricing equation:

$$\mathbb{m}_{\boldsymbol{\pi}_t} := \{\boldsymbol{m}_t : \mathrm{E}_{\boldsymbol{\pi}_t}[\boldsymbol{r}_t \boldsymbol{m}_t] = \mathbf{1}_2\}.$$

Here, we denote the above SDF set as the one-period SDF set with $\boldsymbol{m}_t$ being the SDF that discount the asset payoff at time $kh$ to its price at time $(k-1)h$ for $k = 1,2$, and the multiperiod SDF at $t$ is defined as $\boldsymbol{M}_t = \Pi_{k=1}^{t/h} \boldsymbol{m}_{kh}$ (Cochrane, 2009), which prices a $k$-period payoff and satisfies the following equation

$$\mathrm{E}_{\boldsymbol{\pi}_{kh}}[\boldsymbol{r}_{kh} \boldsymbol{M}_{kh}] = \boldsymbol{M}_{(k-1)h}.$$

Then, the multiperiod SDF set at $t$ can be written in the form $\mathbb{M}_{\boldsymbol{\pi}_t} = \Pi_{k=1}^{t/h} \mathbb{m}_{\boldsymbol{\pi}_{kh}}$.

To examine the evolution of market incompleteness over time, we focus on the single-period SDF set at time $t$ such that for every $\boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t$

$$\mathbb{m}_{\boldsymbol{\pi}_t} = \left\{ \boldsymbol{m}_t \in \mathbb{R}^3_* : \begin{bmatrix} m_t^1 \\ m_t^2 \\ m_t^3 \end{bmatrix} = \begin{bmatrix} \frac{1 - r_t^{1,2}}{r_t^{1,1} - r_t^{1,2}} \\ \frac{r_t^{1,1} - 1}{r_t^{1,1} - r_t^{1,2}} \\ 0 \end{bmatrix} \boldsymbol{\pi}_t^{-1} + \alpha \begin{bmatrix} \frac{r_t^{1,2} - r_t^{1,3}}{r_t^{1,1} - r_t^{1,2}} \\ -\frac{r_t^{1,1} - r_t^{1,3}}{r_t^{1,1} - r_t^{1,2}} \\ 1 \end{bmatrix} \boldsymbol{\pi}_t^{-1}, \alpha \in \mathbb{R}_* \right\},$$

where $m_t^s \equiv \boldsymbol{m}_t(\omega_t^s)$ and $\alpha = (1 - \pi_t^1 - \pi_t^2)m_t^3$, and the set that contains all single-period SDFs in each period $t$ is $\boldsymbol{c}_t \coloneqq \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\}$.

Further, analogous to the results proved in the 1-1-2-3 case, the following proposition and theorems hold.

**Proposition 3.2.3.** *For $t = kh$, consider the metric space $(\bar{\boldsymbol{\Pi}}_t, d)$ such that*

$$\bar{\boldsymbol{\Pi}}_t = \left\{\boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \pi_t^3\right]^\top : \sum_{s=1}^{3} \pi_t^s = 1, \pi_t^1, \pi_t^2 > 0, \pi_t^3 \geq 0\right\}$$

*and $d$ is the Euclidean distance metric. Then, the set of limit points of $\boldsymbol{\Pi}_t$ in $(\bar{\boldsymbol{\Pi}}_t, d)$ is*

$$L(\boldsymbol{\Pi}_t) = \left\{\boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \pi_t^3\right]^\top : \sum_{s=1}^{3} \pi_t^s = 1, \pi_t^1, \pi_t^2 > 0, \pi_t^3 \geq 0\right\},$$

*and the set of boundary points of $\boldsymbol{\Pi}_t$ in $\bar{\boldsymbol{\Pi}}_t$ is*

$$\partial\boldsymbol{\Pi}_t = \left\{\boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \pi_t^3\right]^\top : \sum_{s=1}^{3} \pi_t^s = 1, \pi_t^1, \pi_t^2 > 0, \pi_t^3 = 0\right\}.$$

**Theorem 3.2.5.** *Consider the metric space $(\bar{\boldsymbol{c}}_t, d_1)$ such that $\bar{\boldsymbol{c}}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \bar{\boldsymbol{\Pi}}_t\}$ and for $\boldsymbol{x}_t, \boldsymbol{y}_t \in \bar{\boldsymbol{\Pi}}_t$,*

$$d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}) = \inf_{\boldsymbol{w}_t}\left\{\sum_{s_{\boldsymbol{x}_t}=1}^{S_{\boldsymbol{x}_t}} \sum_{s_{\boldsymbol{y}_t}=1}^{S_{\boldsymbol{y}_t}} w_t^{s_{\boldsymbol{x}_t} s_{\boldsymbol{y}_t}} x_t^{s_{\boldsymbol{x}_t}} d_2\left(m_t^{s_{\boldsymbol{x}_t}}, m_t^{s_{\boldsymbol{y}_t}}\right) : \boldsymbol{w}_t \in W(\boldsymbol{x}_t, \boldsymbol{y}_t)\right\}, \tag{3.13}$$

*where $W(\boldsymbol{x}_t, \boldsymbol{y}_t) \coloneqq \{\boldsymbol{w}_t \in \mathbb{R}_+^{S_{\boldsymbol{y}_t} \times S_{\boldsymbol{x}_t}} : \boldsymbol{w}_t^\top \mathbf{1}_{S_{\boldsymbol{y}_t}} = \boldsymbol{x}_t, \boldsymbol{w}_t \boldsymbol{x}_t = \boldsymbol{y}_t\}$ is the set of transport plans between $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$. $S_{\boldsymbol{z}_t}$ is the number of states with non-zero probabilities and the subscript $s_{\boldsymbol{z}_t}$ is the index of the elements in the vector under the physical probability $\boldsymbol{z}_t \in \bar{\boldsymbol{\Pi}}_t$. For all $s_{\boldsymbol{x}_t} = 1, 2, \ldots, S_{\boldsymbol{x}_t}$ and $s_{\boldsymbol{y}_t} = 1, 2, \ldots, S_{\boldsymbol{y}_t}$,*

$$d_2\left(m_t^{s_{\boldsymbol{x}_t}}, m_t^{s_{\boldsymbol{y}_t}}\right) = \left|v^{s_{\boldsymbol{x}_t}} - v^{s_{\boldsymbol{y}_t}}\right| + \left|u^{s_{\boldsymbol{x}_t}} - u^{s_{\boldsymbol{y}_t}}\right|,$$

*where*

$$\begin{cases} \boldsymbol{v}(\boldsymbol{x}_t) = \begin{bmatrix} \frac{1-r_t^{1,2}}{r_t^{1,1}-r_t^{1,2}} \\ \frac{r_t^{1,1}-1}{r_t^{1,1}-r_t^{1,2}} \\ 0 \end{bmatrix} \boldsymbol{x}_t^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}_t) = \begin{bmatrix} \frac{r_t^{1,2}-r_t^{1,3}}{r_t^{1,1}-r_t^{1,2}} \\ -\frac{r_t^{1,1}-r_t^{1,3}}{r_t^{1,1}-r_t^{1,2}} \\ 1 \end{bmatrix} \boldsymbol{x}_t^{-1}, & \boldsymbol{x}_t \in \boldsymbol{\Pi}_t; \\ \boldsymbol{v}(\boldsymbol{x}_t) = \begin{bmatrix} \frac{1-r_t^{1,2}}{r_t^{1,1}-r_t^{1,2}} \\ \frac{r_t^{1,1}-1}{r_t^{1,1}-r_t^{1,2}} \end{bmatrix} \boldsymbol{x}_t^{-1} \text{ and } \boldsymbol{u}(\boldsymbol{x}_t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \boldsymbol{x}_t^{-1}, & \boldsymbol{x}_t \in \partial\boldsymbol{\Pi}_t. \end{cases}$$

*Then, the set of limit points of* $\boldsymbol{c}_t$ *in* $\bar{\boldsymbol{c}}_t$ *can be denoted as*

$$L(\boldsymbol{c}_t) = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\} \cup \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t\},$$

*where, for any* $\boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t$,

$$\mathbb{m}_{\boldsymbol{\pi}_t} = \left\{ \begin{bmatrix} m_t^1 \\ m_t^2 \end{bmatrix} = \begin{bmatrix} \frac{1-r_t^{1,2}}{r_t^{1,1}-r_t^{1,2}} \\ \frac{r_t^{1,1}-1}{r_t^{1,1}-r_t^{1,2}} \end{bmatrix} \boldsymbol{\pi}_t^{-1} \right\},$$

*and the set of boundary points of* $\boldsymbol{c}_t$ *is then* $\partial\boldsymbol{c}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t\}$.

**Theorem 3.2.6.** *Let* $\boldsymbol{\Pi}_t$ *be the set of physical probability measures satisfying Equation (3.12). Let* $\mathbb{m}_{\boldsymbol{\pi}_t}$ *be the identified SDF set in the 1-1-3-3 case given the probability measure* $\boldsymbol{\pi}_t$ *in* $\boldsymbol{\Pi}_t$*, and let* $\boldsymbol{c}_t$ *be the combined SDF set such that* $\boldsymbol{c}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\}$. *Then,* $\boldsymbol{c}_t$ *is a convex set.*

*Let* $(\bar{\boldsymbol{c}}_t, d_1)$ *be the metric space such that* $\bar{\boldsymbol{c}}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \bar{\boldsymbol{\Pi}}_t\}$ *and for* $\boldsymbol{x}_t, \boldsymbol{y}_t \in \bar{\boldsymbol{\Pi}}_t$, $d_1$ *is as defined in Equation (3.13). Then,* $\boldsymbol{c}_t$ *is open, bounded and not compact under* $(\bar{\boldsymbol{c}}_t, d_1)$.

Last, we derive the degree of market incompleteness based on the set properties in Theorems 3.2.5 and 3.2.6. Again, since at time 0, the complete and incomplete market SDFs are assumed to be 1, the distance between them is 0 following the metric $d_1$. Thus, given $\{\boldsymbol{x}_{kh} \in \bar{\boldsymbol{\Pi}}_{kh}\}_{k=1}^{t/h}$, the degree of market incompleteness at $t$ is defined as in Equation (3.14), which is the average of the least transport costs from $\mathbb{m}_{\boldsymbol{x}_{kh}} \in \bar{\boldsymbol{c}}_{kh}$ to $\mathbb{m}_{\boldsymbol{y}_{kh}} \in \partial\boldsymbol{c}_{kh}$ from time $h$ to $t$, indicating that we weigh the degree of market incompleteness equally across subperiods.

$$MI\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{t/h}\right) = \frac{h}{t}\sum_{k=1}^{t/h} \min_{\mathbb{m}_{\boldsymbol{y}_{kh}} \in \partial\boldsymbol{c}_{kh}} d_1\left(\mathbb{m}_{\boldsymbol{x}_{kh}}, \mathbb{m}_{\boldsymbol{y}_{kh}}\right). \tag{3.14}$$

As the subperiod degrees of market incompleteness are functions of their subperiod asset returns, which are uncorrelated, we take the average of them so that $MI\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{t/h}\right)$ is not monotonic in $t$. Moreover, since $d_1$ is a valid metric and the complete market SDF is the boundary point of the incomplete market SDF set in each subperiod, the degree of market incompleteness equals zero only when markets are dynamically complete, i.e., when markets are complete at every subperiods. Hence, the estimated $\hat{MI}(\{\boldsymbol{x}_h\})$ and $\hat{MI}\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{2}\right)$ depict the evolution of market incompleteness over the time interval $[0,1]$.

### 3.2.4 Generalization of the Discrete-time Setting

Finally, we consider a generalized discrete-time setting in which there are a finite number of additional states rather than only one, while still allowing for any types of idiosyncratic risks. The setup is formalized as follows.

**Assumption 3.2.2.** *Suppose that there are one risk-free bond and $A \in \mathbb{N}$ risky assets. Let $S \geq A+2$ and $(\Omega, \mathcal{F}, P)$ be the complete probability space, where $\Omega$, $\mathcal{F}$ and $P$ are the same as defined in Assumption 3.2.1. Suppose that there are $K \geq 1$ equally-spaced subperiods in $[0,1]$, and let $h = 1/K$ be the time window. All assets are long-lived and available for trading at time points $\{0, h, \ldots, (K-1)h, 1\}$, and there are $S$ states at each $t = \{kh\}_{k=1,2,\ldots,K}$ such that $\Omega_t = (\omega_t^s)_{s=1,2,\ldots,S}$. Let $P_t(\omega_t^s) = \pi_t^s$ be the physical probability of state $\omega_t^s$, where $\pi_t^s$ are strictly positive scalars for all $s$ in incomplete markets, while $[\boldsymbol{\pi}_t^{\bar{s}}]_{\bar{s}=A+2,A+3,\ldots,S} = \mathbf{0}_{S-A-1}$ when the markets are complete. There exists a set $\mathcal{P}_t$ of complete probability measures at each $t = \{kh\}_{k=1,\ldots,K}$ such that $P_t \in \mathcal{P}_t$.*

Assumption 3.2.2 is a generalization of Assumption 3.2.1, where the amount of unhedgeable risks at time $t$ is no longer restricted to be one, and we allow for multiple additional states in each subperiod. Correspondingly, the set of physical probabilities at $t$ when the markets are incomplete is[11]

$$\mathbf{\Pi}_t = \left\{ \left[\pi_t^1, \pi_t^2, \ldots, \pi_t^S\right] \in \mathbb{R}_{++}^S : \sum_{s=1}^{S} \pi_t^s = 1 \right\}. \tag{3.15}$$

At time $t = kh$, the gross rate of return vector of length $A+1$ realized at state $s$ is $\boldsymbol{r}_t(\omega_t^s) = \left[r_t^{1,s}, r_t^{2,s}, \ldots, r_t^{A,s}, r_t^0\right]^\top$, where $r_t^{a,s}$ denotes the return of the risky asset $a$ in state $s$, and $r_t^0$ denotes the risk-free rate. Let $\boldsymbol{r}_t = \left[\boldsymbol{r}_t(\omega_t^s)\right]_{s=1,2,\ldots,S}$, when the assumption that the second-moment of $\boldsymbol{r}_t$ is nonsingular holds, at the end of each subperiod $t$, we have the random variable $\boldsymbol{m}_t : \Omega_t \to \mathbb{m}_{\boldsymbol{\pi}_t}$, where $\mathbb{m}_{\boldsymbol{\pi}_t}$ is the set of SDFs under $P_t \in \mathcal{P}_t$ that satisfies the asset pricing equation:

$$\mathbb{m}_{\boldsymbol{\pi}_t} := \left\{ \boldsymbol{m}_t : \mathrm{E}_{\boldsymbol{\pi}_t}[\boldsymbol{r}_t \boldsymbol{m}_t] = \mathbf{1}_{A+1}. \right\}. \tag{3.16}$$

Subsequently, let $\boldsymbol{\alpha}_t = [\pi_t^{\bar{s}} m_t^{\bar{s}}]_{\bar{s}=A+2,A+3,\ldots,S} \in \mathbb{R}_*^{S-A-1}$ be the vector of free variables, we can derive the single-period SDF set at time $t$ in the form such that for every $\boldsymbol{\pi}_t \in \mathbf{\Pi}_t$

$$\mathbb{m}_{\boldsymbol{\pi}_t} = \left\{ \boldsymbol{m}_t \in \mathbb{R}_*^S : \boldsymbol{m}_t = \boldsymbol{v}_t(\boldsymbol{\pi}_t) + \boldsymbol{u}_t(\boldsymbol{\pi}_t)\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_t \in \mathbb{R}_*^{S-A-1} \right\},$$

[11] Here, we restrict all $\pi_t^s$'s to be positive under incomplete markets, instead of letting $\pi_t^s \geq 0$ for $s \geq A+2$, because the latter can be simplified to a lower-dimensional case. For instance, if $\pi_t^S = 0$ under both complete and incomplete markets, then, our setup can be reduced to an $(S-1)$-dimensional case.

where

$$\boldsymbol{v}_t(\boldsymbol{\pi}_t) = \begin{bmatrix} (\boldsymbol{r}'_t)^{-1}\,\mathbf{1}_{A+1} \\ \mathbf{0}_{S-A-1} \end{bmatrix} \boldsymbol{\pi}_t^{-1} \text{ and } \boldsymbol{u}_t(\boldsymbol{\pi}_t) = \begin{bmatrix} -\,(\boldsymbol{r}'_t)^{-1}\,(\boldsymbol{r}''_t) \\ \mathbf{1}_{S-A-1} \end{bmatrix} \boldsymbol{\pi}_t^{-1}$$

with $\boldsymbol{r}'_t = (\boldsymbol{r}_t)_{*(1:A+1)}$ and $\boldsymbol{r}''_t = (\boldsymbol{r}_t)_{*(A+2:S)}$. Lastly, the combined set $\boldsymbol{c}_t$ of $\mathbb{m}_{\boldsymbol{\pi}_t}$'s for all $\boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t$ is defined as $\boldsymbol{c}_t := \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\}$.

Now, we demonstrate that the results in previous special cases hold in the generalized setting. The following proposition indicates that the probability distribution under complete markets is the boundary point of the set of probabilities under incomplete markets, and its proof is discussed in Appendix C.10.

**Proposition 3.2.4.** *Consider the metric space $(\bar{\boldsymbol{\Pi}}_t, d)$ such that*

$$\bar{\boldsymbol{\Pi}}_t = \boldsymbol{\Pi}_t \cup \Big\{ \boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \ldots, \pi_t^S\right]^\top : \sum_{s=1}^{S} \pi_t^s = 1, \pi_t^s > 0 \textit{ for } s = 1, 2, \ldots, A+1, \\ \pi_t^s = 0 \textit{ for } s = A+2, A+3\ldots, S \Big\}$$

*and $d$ is the Euclidean distance metric.*

*Then, the set of limit points of $\boldsymbol{\Pi}_t$ in $(\bar{\boldsymbol{\Pi}}_t, d)$ is $L(\boldsymbol{\Pi}_t) = \bar{\boldsymbol{\Pi}}_t$, and the set of boundary points of $\boldsymbol{\Pi}_t$ in $\bar{\boldsymbol{\Pi}}_t$ is*

$$\partial\boldsymbol{\Pi}_t = \Big\{ \boldsymbol{\pi}_t = \left[\pi_t^1, \pi_t^2, \ldots, \pi_t^S\right]^\top : \sum_{s=1}^{S} \pi_t^s = 1, \pi_t^s > 0 \textit{ for } s = 1, 2, \ldots, A+1, \\ \pi_t^s = 0 \textit{ for } s = A+2, A+3\ldots, S \Big\}.$$

Then, we can establish the following result such that the constructed incomplete SDF set has its boundary point to be the complete market SDF under the defined metric space. The proof of this result is presented in Appendix C.11

**Theorem 3.2.7.** *Consider the metric space $(\bar{\boldsymbol{c}}_t, d_1)$ with $\bar{\boldsymbol{c}}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi_t} \in \bar{\boldsymbol{\Pi}}_t\}$ and $d_1$ is as defined in Equation (3.13). Then, the set of limit points of $\boldsymbol{c}_t$ in $(\bar{\boldsymbol{c}}_t, d_1)$ is*

$$L(\boldsymbol{c}_t) = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\} \cup \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t\},$$

*where, for any $\boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t$,*

$$\mathbb{m}_{\boldsymbol{\pi}_t} = \left\{ \boldsymbol{m}_t = \left[ (\boldsymbol{r}'_t)^{-1}\,\mathbf{1}_{A+1}(\boldsymbol{\pi}_t)^{-1}_{1:(A+1)} \right] \right\},$$

*and the set of boundary points of $\boldsymbol{c}_t$ is then $\partial\boldsymbol{c}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \partial\boldsymbol{\Pi}_t\}$.*

The following lemma, derived from Theorem 3.2.7, implies that for every incomplete market SDF set, $\mathbb{m}_{\boldsymbol{x}_t} \in \boldsymbol{c}_t$, there exists a complete market SDF, $\mathbb{m}_{\boldsymbol{y}_t} \in \partial \boldsymbol{c}_t$ that minimizes the distance between them. This lemma enables us to further explore the set properties of the incomplete market SDF as well as the degree of market incompleteness. Its proof is presented in Appendix C.12.

**Lemma 3.2.3.** *For every $\mathbb{m}_{\boldsymbol{x}_t} \in \boldsymbol{c}_t$, there exists $\mathbb{m}_{\boldsymbol{y}_t^*}$ such that*

$$\mathbb{m}_{\boldsymbol{y}_t^*} = \underset{\mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t}{\arg\min} \left\{ d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}) \right\}.$$

The next result establishes the *convexity*, *openness*, *boundedness*, and *non-compactness*, for $\boldsymbol{c}_t$ in the generalized case with its proof discussed in Appendix C.13.

**Theorem 3.2.8.** *Let $\boldsymbol{\Pi}_t$ be the set of all the probability density measures under $P_t$ satisfying Equation (3.15).*

*Let $\mathbb{m}_{\boldsymbol{\pi}_t}$ be the identified SDF set in the generalized case satisfying Equation (3.16) given $\boldsymbol{\pi}_t$ in $\boldsymbol{\Pi}_t$.*

*Let $\boldsymbol{c}_t$ be the combined SDF set such that $\boldsymbol{c}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_t\}$. Then, $\boldsymbol{c}_t$ is a convex set.*

*Let $(\bar{\boldsymbol{c}}_t, d_1)$ be the metric space such that $\bar{\boldsymbol{c}}_t = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \bar{\boldsymbol{\Pi}}_t\}$ and for $\boldsymbol{x}_t, \boldsymbol{y}_t \in \bar{\boldsymbol{\Pi}}_t$, $d_1$ is as defined in Equation (3.13). Then, $\boldsymbol{c}_t$ is open, bounded and not compact under $(\bar{\boldsymbol{c}}_t, d_1)$.*

Last, based on the set properties in Theorems 3.2.7 and 3.2.8, given $\left\{\boldsymbol{x}_{kh} \in \bar{\boldsymbol{\Pi}}_{kh}\right\}_{k=1}^{t/h}$, the degree of market incompleteness at $t$ is defined as in Equation (3.17), which is the mean of the least transport costs from $\mathbb{m}_{\boldsymbol{x}_{kh}} \in \bar{\boldsymbol{c}}_{kh}$ to $\mathbb{m}_{\boldsymbol{y}_{kh}} \in \partial \boldsymbol{c}_{kh}$ from time $h$ up to $t$[12]:

$$MI\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{t/h}\right) = \frac{h}{t} \sum_{k=1}^{t/h} \min_{\mathbb{m}_{\boldsymbol{y}_{kh}} \in \partial \boldsymbol{c}_{kh}} d_1\left(\mathbb{m}_{\boldsymbol{x}_{kh}}, \mathbb{m}_{\boldsymbol{y}_{kh}}\right). \tag{3.17}$$

The subperiod degrees of market incompleteness are functions of their subperiod asset returns, which are uncorrelated, then by taking the average of these subperiod transport costs, we get the degree of market incompleteness at $t$, $MI\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{t/h}\right)$, which is not monotonic in $t$. Moreover, as $d_1$ is a valid metric and the complete market SDF is the boundary point of the incomplete market SDF set in each subperiod, the degree of market incompleteness equals zero only when the markets are dynamically complete, i.e., when the markets are complete at every subperiods. Hence, for $t \in (0,1]$, the estimated $\hat{MI}\left(\{\boldsymbol{x}_{kh}\}_{k=1}^{t/h}\right)$'s depict the evolution of market incompleteness over time.

[12]Again, we assume that the complete and incomplete market SDFs are 1 at time 0, and the distance between them is 0 following the metric $d_1$.

## 3.3 Continuous-time Setting

The modelling of the SDF set and the degree of market incompleteness in the continuous-time setting is similar to that used in its discrete-time counterpart, but there are differences. Particularly, in order to implement our approach in empirical works, we further parameterize the market incompleteness by specifying that the asset prices are generated by the jump diffusion processes, which constitute an important class of incomplete market models and are realistic in practice (Kaido and White, 2009). We formalize our setup as follows

**Assumption 3.3.1.** *Suppose that there are one risk-free bond and $A \in \mathbb{N}$ risky assets in the market. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,1]}, P)$ be the complete filtered probability space, where $\Omega$, $\mathcal{F}$ and $P$ are the same as defined in Assumption 3.2.2, and the filtration $\{\mathcal{F}_t\} = \{\mathcal{F}_t\}_{t\in[0,1]}$ is assumed to satisfy the usual properties (Protter, 2005). There exists a set $\mathcal{P}$ of complete probability measures on $(\Omega, \mathcal{F})$ such that $P \in \mathcal{P}$.*

In incomplete markets, we have $\mathbb{R}^A$ –valued risky asset price process $\{\boldsymbol{S}_t\}$, which solves the stochastic differential equation (SDE)

$$\frac{d\boldsymbol{S}_t}{\boldsymbol{S}_{t-}} = \boldsymbol{\mu}_t^B dt + \boldsymbol{\sigma}_t^B d\boldsymbol{B}_t + \boldsymbol{J}_t d\tilde{\boldsymbol{N}}_t, \tag{3.18}$$

where $\{\boldsymbol{\mu}_t^B\}$ is an $\mathbb{R}^A$-valued adapted drift process, $\{\boldsymbol{\sigma}_t^B\}$ is an $\mathbb{R}^{A\times A}$-valued adapted diffusion coefficient process. $\boldsymbol{J}_t$ is a random jump amplitude, which is predictable and $\boldsymbol{J}_t > -1$, implying that all elements in $\boldsymbol{S}_t$ remain positive, consistent with the limited liability provision (Aït-Sahalia et al., 2009). Then, it is convenient to have $\boldsymbol{J}_t = \exp(\boldsymbol{Q}_t) - \boldsymbol{1}_A$ as in Hanson and Westman (2002), where $\boldsymbol{Q}_t$ follows a normal distribution with mean $\boldsymbol{\mu}_t^J$ and standard deviation $\boldsymbol{\sigma}_t^J$. $\{\boldsymbol{B}_t\}$ is a vector of $A$ independent Brownian motions under $P$ and $\tilde{\boldsymbol{N}}_t = \boldsymbol{N}_t - \boldsymbol{v}_t(dx)t$ is the compensated martingales of Poisson process $\boldsymbol{N}_t$ with mean measure $\boldsymbol{v}_t(dx)t$, where $\boldsymbol{v}_t(dx) \geq 0$ is taken to be the Lévy measure associated with an $A$-dimensional pure-jump Lévy process. Thus, $\boldsymbol{v}_t(dx)$ is a $\sigma$-finite measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$[13] such that $\boldsymbol{v}_t(\{0\}) = 0$, suggesting that $\boldsymbol{v}$ does not have mass on 0, and

$$\int_{\mathbb{R}} \min\left(1, |x|^2\right) \boldsymbol{v}_t(dx) < \infty, \tag{3.19}$$

so the jumps have finite variation. $\{\boldsymbol{B}_t\}$ and $\tilde{\boldsymbol{N}}_t$ are independent under $P$ and are adapted to $\{\mathcal{F}_t\}$. We require that

$$\mathbb{P}\left[\int_0^t \left(\left|\boldsymbol{\mu}_s^B\right| + {\boldsymbol{\sigma}_s^B}^2 + \boldsymbol{J}_s^2 \boldsymbol{v}_s(dx)\right) ds < \infty\right] = 1 \tag{3.20}$$

[13] We use $\mathcal{B}(\mathbb{R})$ to denote the Borel $\sigma$-algebra.

for $t \geq 0$, which is a sufficient restriction to ensure that the integral with respect to the compensated Poisson random measure exists for both small and large jumps. We assume that the market is built with a risk-free bond with a known rate of return of $r_t$.

Given the $\mathbb{R}^A$-valued adapted processes $\{\boldsymbol{\psi}_t\}_{t\geq 0}$ and $\{\boldsymbol{\gamma}_t\}_{t\geq 0}$, the Girsanov transformation defines the new adapted processes $\{\bar{\boldsymbol{B}}_t\}$ and $\{\bar{\boldsymbol{N}}_t\}$ by adjusting the original Brownian motion and the compensated martingales of Poisson process:

$$\bar{\boldsymbol{B}}_t = \boldsymbol{B}_t + \int_0^t \boldsymbol{\psi}_s ds \text{ and } \bar{\boldsymbol{N}}_t = \tilde{\boldsymbol{N}}_t + \int_0^t \boldsymbol{v}_s(dx)\boldsymbol{\gamma}_s ds.$$

Then, the asset return process under the risk-neutral probability measure can be written as

$$\frac{d\boldsymbol{S}_t}{\boldsymbol{S}_{t-}} = r_t \mathbf{1}_A^B dt + \boldsymbol{\sigma}_t^B d\bar{\boldsymbol{B}}_t + \boldsymbol{J}_t d\bar{\boldsymbol{N}}_t$$

and the existence of the SDF holds only for $(\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t)$ satisfying the condition below

$$\boldsymbol{\mu}_t^B - r_t \mathbf{1}_A - \boldsymbol{\sigma}_t^B \boldsymbol{\psi}_t - \boldsymbol{J}_t \boldsymbol{\gamma}_t \boldsymbol{v}_t(dx) = 0, \ a.s. - P.$$

Such vectors are called the market prices of risk, where $\{\boldsymbol{\psi}_t\}_{t\geq 0}$ is the adapted Brownian market price of risk and $\{\boldsymbol{\gamma}_t\}_{t\geq 0}$ is the predictable jump market price of risk, and $\boldsymbol{\gamma}_t < 1$ for $t \geq 0$.

Let $\boldsymbol{\phi}_t = (\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)) \in \boldsymbol{\Phi}_t$, where

$$\boldsymbol{\Phi}_t = \left\{ \left( \boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx) \right) : \boldsymbol{v}_t(dx) > \mathbf{0}_A \right\}$$

is an admissible parameter space under $P \in \mathcal{P}$. When markets are incomplete, the market prices of risk form the set

$$\boldsymbol{\Gamma}(\boldsymbol{\phi}_t) = \left\{ (\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t) : \boldsymbol{\mu}_t^B - r_t \mathbf{1}_A - \boldsymbol{\sigma}_t^B \boldsymbol{\psi}_t - \boldsymbol{J}_t \boldsymbol{\gamma}_t \boldsymbol{v}_t(dx) = 0 \right\}. \tag{3.21}$$

Let $\boldsymbol{\alpha}_t = \ln\left[ (\mathbf{1}_A - \boldsymbol{\gamma}_t)^{-1} \right]$, Equation (3.21) is transformed to

$$\begin{aligned}\boldsymbol{\Gamma}(\boldsymbol{\phi}_t) = \Big\{ (\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t) : \boldsymbol{\psi}_t = (\boldsymbol{\sigma}_t^B)^{-1} \left( \boldsymbol{\mu}_t^B - r_t \mathbf{1}_A \right) - \left( \boldsymbol{\sigma}_t^B \right)^{-1} \left( \boldsymbol{J}_t \left( \mathbf{1}_A - e^{-\boldsymbol{\alpha}_t} \right) \boldsymbol{v}_t(dx) \right), \\ \boldsymbol{\gamma}_t = \mathbf{1}_A - e^{-\boldsymbol{\alpha}_t}, \boldsymbol{\alpha}_t \in \mathbb{R}^A \Big\}.\end{aligned} \tag{3.22}$$

Correspondingly, for $(\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t) \in \boldsymbol{\Gamma}(\boldsymbol{\phi}_t)$, the SDF process $\{\boldsymbol{M}(\boldsymbol{\phi}_t)\}_{t\geq 0}$ follows the dynamic form

$$\frac{d\boldsymbol{M}(\boldsymbol{\phi}_t)}{\boldsymbol{M}(\boldsymbol{\phi}_{t-})} = -\left[ r_t \mathbf{1}_A dt + \boldsymbol{\psi}_t d\boldsymbol{B}_t + \boldsymbol{\gamma}_t d\tilde{\boldsymbol{N}}_t \right]$$

with the solution

$$\boldsymbol{M}(\boldsymbol{\phi}_t) = \exp\left(-\int_0^t r_s \mathbf{1}_A ds - \int_0^t \boldsymbol{\psi}_s d\boldsymbol{B}_s - \frac{1}{2}\int_0^t \boldsymbol{\psi}_s^2 ds\right) \times \exp\left(-\int_0^t \boldsymbol{\alpha}_s d\tilde{\boldsymbol{N}}_s - \int_0^t \left(e^{-\boldsymbol{\alpha}_s} - \mathbf{1}_A + \boldsymbol{\alpha}_s\right) \boldsymbol{v}_t(dx) ds\right).$$

We shall restrict $\boldsymbol{M}(\boldsymbol{\phi}_t)$ to be a $P$-square integrable martingale over the time interval $[0,1]$, i.e., $\sup_{t\in[0,1]} \mathrm{E}\left[\boldsymbol{M}^2(\boldsymbol{\phi}_t)\right] < \infty$. Then, the SDF set is

$$\mathbb{M}(\boldsymbol{\phi}_t) = \Bigg\{\boldsymbol{M}(\boldsymbol{\phi}_t) = \exp\left(-\int_0^t r_s \mathbf{1}_A ds - \int_0^t \boldsymbol{\psi}_s d\boldsymbol{B}_s - \frac{1}{2}\int_0^t \boldsymbol{\psi}_s^2 ds\right) \times \exp\left(-\int_0^t \boldsymbol{\alpha}_s d\tilde{\boldsymbol{N}}_s - \int_0^t \left(e^{-\boldsymbol{\alpha}_s} - \mathbf{1}_A + \boldsymbol{\alpha}_s\right) \boldsymbol{v}_t(dx) ds\right), \boldsymbol{\alpha}_t = \ln\left[(\mathbf{1}_A - \boldsymbol{\gamma}_t)^{-1}\right], (\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t) \in \boldsymbol{\Gamma}(\boldsymbol{\phi}_t)\Bigg\},$$

and the set that contains all SDFs under $P \in \mathcal{P}$ is defined as $\boldsymbol{C}_t := \{\mathbb{M}_t(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_t\}$. Analogous to the discrete case, in order to analyze the evolution of the degree of market incompleteness, we frame the following discussion in terms of the SDF set including all possible SDFs that price payoff over an infinitesimal time interval $[t, t+dt)$:

$$\begin{aligned} \mathbb{m}(\boldsymbol{\phi}_t) &= \{\boldsymbol{m}(\boldsymbol{\phi}_t) = \boldsymbol{M}(\boldsymbol{\phi}_t)/\boldsymbol{M}(\boldsymbol{\phi}_{t-}), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_t\} \\ &= \Bigg\{\boldsymbol{m}(\boldsymbol{\phi}_t) = \exp\left(-r_t \mathbf{1}_A dt - \boldsymbol{\psi}_t d\boldsymbol{B}_t - \frac{1}{2}\boldsymbol{\psi}_t^2 dt\right) \times \exp\left(-\boldsymbol{\alpha}_t d\tilde{\boldsymbol{N}}_t - \left(e^{-\boldsymbol{\alpha}_t} - \mathbf{1}_A + \boldsymbol{\alpha}_t\right) \boldsymbol{v}_t(dx) dt\right), \\ &\qquad \boldsymbol{\alpha}_t = \ln\left[(\mathbf{1}_A - \boldsymbol{\gamma}_t)^{-1}\right], (\boldsymbol{\psi}_t, \boldsymbol{\gamma}_t) \in \boldsymbol{\Gamma}(\boldsymbol{\phi}_t)\Bigg\}, \end{aligned} \tag{3.23}$$

and thus, this SDF discounts the payoff at $t + dt$ to its price at $t$. Accordingly, the set that contains all $\mathbb{m}(\boldsymbol{\phi}_t)$ for $P \in \mathcal{P}$ is $\boldsymbol{c}_t := \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_t\}$.

### 3.3.1 Set Properties of $\boldsymbol{C}_t$

Similar to the discrete-time framework, we first verify that the boundary point of the proposed SDF set is indeed the one under the complete market. Let $\bar{\boldsymbol{\Phi}}_t := \{(\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)) : \boldsymbol{v}_t(dx) \geq \mathbf{0}_A\}$ be the admissible parameter space, the following proposition establishes the limit and boundary points of $\boldsymbol{\Phi}_t$ under $\bar{\boldsymbol{\Phi}}_t$. The proof is shown in Appendix C.14.

**Proposition 3.3.1.** *Consider the metric space* $(\bar{\mathbf{\Phi}}_t, d)$ *such that*

$$\bar{\mathbf{\Phi}}_t := \left\{\left(\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)\right) : \boldsymbol{v}_t(dx) \geq \mathbf{0}_A\right\}$$

*and* $d$ *is the Euclidean norm. Then, the set of limit points of* $\mathbf{\Phi}_t$ *in* $(\bar{\mathbf{\Phi}}_t, d)$ *is*

$$L(\mathbf{\Phi}_t) = \left\{\left(\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)\right) : \boldsymbol{v}_t(dx) \geq \mathbf{0}_A\right\},$$

*and the set of boundary points of* $\mathbf{\Phi}_t$ *in* $(\bar{\mathbf{\Phi}}_t, d)$ *is*

$$\partial\mathbf{\Phi}_t = \left\{\left(\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)\right) : \boldsymbol{v}_t(dx) = \mathbf{0}_A\right\}.$$

The next result indicates that with the continuous-time setup, the complete market SDF is indeed the boundary point of the incomplete market SDF set. The proof is presented in Appendix C.15.

**Theorem 3.3.1.** *Consider the metric space* $(\bar{\boldsymbol{c}}_t, d_3)$ *such that* $\bar{\boldsymbol{c}}_t = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \bar{\mathbf{\Phi}}_t\}$, *and for* $\boldsymbol{\phi}_t, \boldsymbol{\phi}_t' \in \bar{\mathbf{\Phi}}_t$ *satisfies Equation (3.20), let* $P(\boldsymbol{\phi}_t)$ *and* $P(\boldsymbol{\phi}_t')$ *denote the physical probability measures in* $\mathcal{P}$,

$$d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')) = \inf_{w_t}\left\{\int d_4(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t'))dw_t : w_t \in W(P(\boldsymbol{\phi}_t), P(\boldsymbol{\phi}_t'))\right\}, \tag{3.24}$$

*where* $W(P(\boldsymbol{\phi}_t), P(\boldsymbol{\phi}_t')) := \{w_t : \int w_t dP(\boldsymbol{\phi}_t') = P(\boldsymbol{\phi}_t), \int w_t dP(\boldsymbol{\phi}_t) = P(\boldsymbol{\phi}_t')\}$ *is the set of transport plans between* $P(\boldsymbol{\phi}_t)$ *and* $P(\boldsymbol{\phi}_t')$, *and*

$$d_4\left(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')\right) = |f(\boldsymbol{\phi}_t) - f(\boldsymbol{\phi}_t')|$$

*with*

$$f(\boldsymbol{\phi}_t) = \exp\left(-r_t\mathbf{1}_A dt - g(\boldsymbol{\phi}_t)d\boldsymbol{B}_t - \frac{1}{2}g(\boldsymbol{\phi}_t)^2 dt\right) \times \exp\left(-d\tilde{\boldsymbol{N}}_t - e^{-1}\boldsymbol{v}_t(dx)dt\right),$$

$$g(\boldsymbol{\phi}_t) = \left(\boldsymbol{\sigma}_t^B\right)^{-1}\left(\boldsymbol{\mu}_t^B - r_t\mathbf{1}_A\right) - \left(\boldsymbol{\sigma}_t^B\right)^{-1}\left(\boldsymbol{J}_t\left(\mathbf{1}_A - e^{-\mathbf{1}_A}\right)\boldsymbol{v}_t(dx)\right),$$

*and* $\boldsymbol{J}_t = \exp(\boldsymbol{Q}_t) - \mathbf{1}_A, \boldsymbol{Q}_t \sim N(\boldsymbol{\mu}_t^J, {\boldsymbol{\sigma}_t^J}^2)$.

*Then, the set of limit points of* $\boldsymbol{c}_t$ *in* $\bar{\boldsymbol{c}}_t$ *can be denoted as* $L(\boldsymbol{c}_t) = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \bar{\mathbf{\Phi}}_t\}$ *and the set of boundary points of* $\boldsymbol{c}_t$ *is then* $\partial\boldsymbol{c}_t = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \partial\mathbf{\Phi}_t\}$, *where for any* $\boldsymbol{\phi}_t \in \partial\mathbf{\Phi}_t$,

$$\mathbb{m}(\boldsymbol{\phi}_t) = \left\{\boldsymbol{m}(\boldsymbol{\phi}_t) = \exp\left(-r_t\mathbf{1}_A dt - \boldsymbol{\psi}_t d\boldsymbol{B}_t - \frac{1}{2}\boldsymbol{\psi}_t^2 dt\right), \boldsymbol{\mu}_t^B - r_t\mathbf{1}_A - \boldsymbol{\sigma}_t^B\boldsymbol{\psi}_t = 0\right\}. \tag{3.25}$$

Based on Theorem 3.3.1, we derive the following lemma, which will later be incorporated in the discussion of set properties and the degree of market incompleteness. The proof of Lemma 3.3.1 is presented in Appendix C.16.

**Lemma 3.3.1.** *For every $\mathbb{m}(\boldsymbol{\phi}_t) \in \boldsymbol{c}_t$, there exists $\mathbb{m}(\boldsymbol{\phi}_t^*)$ such that*

$$\mathbb{m}(\boldsymbol{\phi}_t^*) = \underset{\mathbb{m}(\boldsymbol{\phi}_t') \in \partial \boldsymbol{c}_t}{\arg\min} \left\{ d_3 \left( \mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t') \right) \right\}.$$

The next theorem establishes the properties of the incomplete market SDF set, and the proof is provided in Appendix C.17.

**Theorem 3.3.2.** *Let $\mathbb{m}(\boldsymbol{\phi}_t)$ be the identified SDF set given that $\boldsymbol{\phi}_t \in \boldsymbol{\Phi}_t$, and let $\boldsymbol{c}_t$ be the combined SDF set such that $\boldsymbol{c}_t = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_t\}$. Then, $\boldsymbol{c}_t$ is a convex set.*

*Let $(\bar{\boldsymbol{c}}_t, d_3)$ be the metric space such that $\bar{\boldsymbol{c}}_t = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \bar{\boldsymbol{\Phi}}_t\}$, and $d_3$ is as defined in Equation (3.24). Then, $\boldsymbol{c}_t$ is open, bounded and not compact under $(\bar{\boldsymbol{c}}_t, d_3)$.*

### 3.3.2 Measure for Market Incompleteness

Based upon Theorems 3.3.1 and 3.3.2, given $\left\{\boldsymbol{\phi}_i \in \bar{\boldsymbol{\Phi}}_i\right\}_{i \in (0,t]}$, the degree of market incompleteness at $t$ is defined as in Equation (3.26), which is the mean of the least transport cost process from $\mathbb{m}(\boldsymbol{\phi}_i) \in \bar{\boldsymbol{c}}_i$ to $\mathbb{m}(\boldsymbol{\phi}_i') \in \partial \boldsymbol{c}_i$ over $(0, t]$[14]:

$$MI_t \left( \{\boldsymbol{\phi}_i\}_{i \in (0,t]} \right) = \mathrm{E}_t \left[ \min_{\boldsymbol{\phi}_i' \in \partial \boldsymbol{\Phi}_i} d_3 \left( \mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}_i') \right) \right]. \tag{3.26}$$

Since $d_3$ is a valid metric and the complete market SDF is the boundary point of the incomplete market SDF set, the degree of market incompleteness equals zero only when the market is dynamically complete, i.e., the distance between complete and incomplete market SDF sets measured by $d_3$ reduces to zero at every $i$ over the time period $(0, t]$.

The following pointwise properties of $MI(\cdot)$ indicate that $MI\left(\{\boldsymbol{\phi}_i\}_{i \in (0,t]}\right)$ is continuous and not monotone in $t$, which enable us to implement our theoretical results in empirical studies and examine the evolution of market incompleteness over time. The proof is presented in Appendix C.18.

**Theorem 3.3.3.** *The degree of market incompleteness $MI\left(\{\boldsymbol{\phi}_i\}_{i \in (0,t]}\right)$ is continuous on the time interval $(0, 1]$ and is not monotone in $t$.*

[14] Same as in the discrete setting, we assume that the complete and incomplete market SDFs are 1 at time 0, and the distance between them is 0 following the metric $d_3$.

## 3.4 Application

This section illustrates the degree of market incompleteness estimation with four countries' major stock market index composites. We first present the layout of a simple but important special case of our continuous-time setup, which will be used for the demonstration of our market incompleteness measure. Then, we describe the data in Section 3.4.1 and the parameter estimations in Section 3.4.2.

Throughout this section, we consider a running example as follows.

**Assumption 3.4.1.** *Let* $\boldsymbol{R}_t := \ln \boldsymbol{S}_t - \ln \boldsymbol{S}_{t-}$ *be a vector of* $A \in \mathbb{N}$ *log-returns observed at* $t \in [0, 1]$. *Suppose the degree of market incompleteness is evaluated at* $K$ *equally-spaced time points* $\{kh\}_{k=1,\dots,K}$ *with the time window* $h = 1/K$. *To simplify the notation, we use the subscript* $k$ *to denote the parameter that characterizes the return in the time period* $[(k-1)h, kh]$.
*When markets are incomplete, let* $\{\boldsymbol{B}_t\}$ *be a vector of* $A \in \mathbb{N}$ *independent standard Brownian motions under* $P$ *and* $\boldsymbol{N}_t$ *be the Poisson process with mean measure* $\boldsymbol{v}_k(dx)t$, *where* $\boldsymbol{v}_k(dx) \geq 0$ *is taken to be the Lévy measure associated with an* $A$*-dimensional pure-jump Lévy process.* $\{\boldsymbol{B}_t\}$ *and* $\{\boldsymbol{N}_t\}$ *are independent and adapted to the filtration* $\{\mathcal{F}_t\}$. $\boldsymbol{R}_t$ *solves the SDE*[15]

$$\boldsymbol{R}_t = \left(\boldsymbol{\mu}_k^B - \boldsymbol{\sigma}_k^{B^2}/2 - \boldsymbol{v}_k(dx)\boldsymbol{\mu}_k^J\right)dt + \boldsymbol{\sigma}_k^B d\boldsymbol{B}_t + \boldsymbol{Q}_k d\boldsymbol{N}_t,$$

*where* $\boldsymbol{\mu}_k^B \in \mathbb{R}^A$, $\boldsymbol{\sigma}_k^B \in \mathbb{R}^{A\times A}$, $\boldsymbol{Q}_k$ *follows a normal distribution with mean* $\boldsymbol{\mu}_k^J \in \mathbb{R}^A$ *and standard deviation* $\boldsymbol{\sigma}_k^J \in \mathbb{R}^{A\times A}$, *and* $dt$ *is estimated by the observational interval. Moreover, both* $\boldsymbol{\sigma}_k^B$ *and* $\boldsymbol{\sigma}_k^J$ *are diagonal matrices, and the price of the risk-free bond has a known constant rate of return* $r_k$.
*When the markets are complete, let* $\{\boldsymbol{B}_t\}$ *be a vector of* $A \in \mathbb{N}$ *independent standard Brownian motions under* $P$. $\boldsymbol{R}_t$ *solves the SDE*

$$\boldsymbol{R}_t = \left(\boldsymbol{\mu}_k^C - \boldsymbol{\sigma}_k^{C^2}/2\right)dt + \boldsymbol{\sigma}_k^C d\boldsymbol{B}_t,$$

*where* $\boldsymbol{\mu}_k^C \in \mathbb{R}^A$, $\boldsymbol{\sigma}_k^C \in \mathbb{R}^{A\times A}$.

Assumption 3.4.1 ensures that the market prices of risk always lie in a nonrandom time-invariant set over a given time period $[(k-1)h, kh]$. Specifically, for $t \in [(k-1)h, kh]$,

[15]Given that the value of interest is usually the log-return on asset, we transform Equation (3.18) using the stochastic chain rule for Markov processes in continuous time, the detailed derivation can be found in Kushner (1967) and Gihman and Skorohod (2012).

$k = 1, \dots, K$, and $\boldsymbol{\alpha}_k = \ln\left[(\mathbf{1}_A - \boldsymbol{\gamma}_k)^{-1}\right]$, Equation (3.22) becomes

$$\Gamma(\boldsymbol{\phi}_k) = \Big\{(\boldsymbol{\psi}, \boldsymbol{\gamma}) : \boldsymbol{\psi}_k = \left(\boldsymbol{\sigma}_k^B\right)^{-1}\left(\boldsymbol{\mu}_k^B - r_k \mathbf{1}_A\right) - \left(\boldsymbol{\sigma}_k^B\right)^{-1}\left(\boldsymbol{J}_k\left(\mathbf{1}_A - e^{-\boldsymbol{\alpha}_k}\right)\boldsymbol{v}_k(dx)\right), \quad \boldsymbol{\gamma}_k = \mathbf{1}_A - e^{-\boldsymbol{\alpha}_k}, \boldsymbol{\alpha}_k \in \mathbb{R}^A\Big\}.$$

Then, under incomplete markets, the SDF set in Equation (3.23) can be written as

$$\mathbb{m}(\boldsymbol{\phi}_k) = \Bigg\{\boldsymbol{m}(\boldsymbol{\phi}_k) = \exp\left(-r_k \mathbf{1}_A dt - \boldsymbol{\psi}_k d\boldsymbol{B}_t - \frac{1}{2}\boldsymbol{\psi}_k^2 dt\right) \times \exp\left(-\boldsymbol{\alpha}_k d\tilde{\boldsymbol{N}}_t - \left(e^{-\boldsymbol{\alpha}_k} - \mathbf{1}_A + \boldsymbol{\alpha}_k\right)\boldsymbol{v}_k(dx)dt\right), (\boldsymbol{\psi}_k, \boldsymbol{\gamma}_k) \in \boldsymbol{\Gamma}(\boldsymbol{\phi}_k)\Bigg\}.$$

Under complete markets, the SDF set in Equation (3.25) can be written as

$$\mathbb{m}\left(\boldsymbol{\phi}_k^C\right) = \left\{\boldsymbol{m}\left(\boldsymbol{\phi}_k^C\right) = \exp\left(-r_k \mathbf{1}_A dt - \boldsymbol{\psi}_k^C d\boldsymbol{B}_t - \frac{1}{2}\boldsymbol{\psi}_k^{C^2} dt\right), \boldsymbol{\mu}_k^C - r_k \mathbf{1}_A - \boldsymbol{\sigma}_k^C \boldsymbol{\psi}_k = 0\right\}.$$

Hence, given $\{\boldsymbol{\phi}_i \in \boldsymbol{\Phi}_i\}_{i=1}^k$, the degree of market incompleteness at $kh$ is

$$MI_{kh}\left(\{\boldsymbol{\phi}_i\}_{i=1}^k\right) = \frac{1}{k}\sum_{i=1}^k \min_{\boldsymbol{\phi}_i^C \in \partial \boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}_i^C)\right),$$

where

$$d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}_i^C)\right) = \inf_{w_i}\left\{\int d_4\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}_i^C)\right) dw_i : w_i \in W\left(P(\boldsymbol{\phi}_i), P(\boldsymbol{\phi}_i^C)\right)\right\}.$$

$W\left(P(\boldsymbol{\phi}_i), P(\boldsymbol{\phi}_i^C)\right) := \left\{w_i : \int w_i dP(\boldsymbol{\phi}_i^C) = P(\boldsymbol{\phi}_i), \int w_i dP(\boldsymbol{\phi}_i) = P(\boldsymbol{\phi}_i^C)\right\}$ is the set of transport plans between $P(\boldsymbol{\phi}_i)$ and $P(\boldsymbol{\phi}_i^C)$, and $d_4\left(\mathbb{M}(\boldsymbol{\phi}_i), \mathbb{M}(\boldsymbol{\phi}_i^C)\right) = \left|f(\boldsymbol{\phi}_i) - f(\boldsymbol{\phi}_i^C)\right|$ with

$$f(\boldsymbol{\phi}_i) = \exp\left(-r_i \mathbf{1}_A dt - g(\boldsymbol{\phi}_i) d\boldsymbol{B}_t - \frac{1}{2} g(\boldsymbol{\phi}_i)^2 dt\right) \times \exp\left(-d\tilde{\boldsymbol{N}}_t - e^{-1}\boldsymbol{v}_i(dx)dt\right),$$

$$g(\boldsymbol{\phi}_i) = (\boldsymbol{\sigma}_i^B)^{-1}(\boldsymbol{\mu}_i^B - r_i \mathbf{1}_A) - (\boldsymbol{\sigma}_i^B)^{-1}\left(\boldsymbol{J}_i(\mathbf{1}_A - e^{-1})\boldsymbol{v}_i(dx)\right),$$

and $\boldsymbol{J}_i = \exp(\boldsymbol{Q}_i) - 1, \boldsymbol{Q}_i \sim N(\boldsymbol{\mu}_i^J, \boldsymbol{\sigma}_i^{J^2})$;

$$f\left(\boldsymbol{\phi}_i^C\right) = \exp\left(-r_i \mathbf{1}_A dt - g\left(\boldsymbol{\phi}_i^C\right) d\boldsymbol{B}_t - \frac{1}{2} g^2\left(\boldsymbol{\phi}_i^C\right) dt\right)$$

and

$$g\left(\boldsymbol{\phi}_i^C\right) = \left(\boldsymbol{\sigma}_i^C\right)^{-1}\left(\boldsymbol{\mu}_i^C - r_i \mathbf{1}_A\right).$$

### 3.4.1 Data Description

Our empirical study analyzes the financial markets of China, Japan, the United Kingdom (UK), and the United States (US) using publicly available data from Yahoo Finance. Due to the availability of data, the Chinese and the US samples begin in 1994, the UK sample begins in 1995, whereas the Japanese sample begins in 1999, and all samples end in 2021. We use the stock data from CSI 300 index for China, Nikkei 225 index for Japan, and FTSE 350 index for the UK and S&P 500 for the US[16].

The stock data is collected on a daily basis, and to examine the evolution of market incompleteness, we divided the full sample into yearly blocks, i.e., for the US market, there are 27 sub-samples, then $K = 27$ and $h = 1/27$. The daily log return (henceforth, the return) is calculated, and assuming 252 trading days per year, $dt$ is estimated by $\Delta = 1/252$. Further, stocks with less than one-month of data are excluded from each subsample in order to eliminate outliers and ensure the reliability of the estimates.

### 3.4.2 Estimation Algorithm

At each subperiod $[(k-1)h, kh]$, we first estimate the parameters $\boldsymbol{\phi}_k$ under incomplete market assumption using the maximum likelihood estimation (MLE) method, and the parameters $\boldsymbol{\phi}_k^C$ under the complete market assumption using the analytical closed-form expression. To the best of our knowledge, there are not yet an analytic expression of the optimal parameter values for jump diffusion models, and thus, we employ the MATLAB function $fminsearchbnd$, which is developed based upon $fminsearch$ to find the minimum value of the constrained multivariable function using derivative-free method for our estimation. As a prerequisite to applying the $fminsearchbnd$ method, we must first establish an initial estimation of the parameters based on the empirical data. Consistent with Merton (1976)'s definition, in this study, we say that there is a jump in the process when the absolute value of return exceeds some threshold $\epsilon > 0$, which is determined as the minimum absolute value of the 5% and 95% quantiles of returns[17], and then, we divide

[16] The CSI 300 is a capitalization-weighted index that replicates the performance of the top 300 stocks traded on the Shanghai Stock Exchange and the Shenzhen Stock Exchange. The Nikkei 225 index measures the performance of 225 large, publicly owned companies in Japan that span a wide range of industry sectors. The FTSE 350 is a capitalization-weighted index composed of the 350 largest companies listed on the London Stock Exchange. The S&P 500 index is a capitalization-weighted index that represents around 80% of the market capitalization of the New York Stock Exchange.

[17] Other quantiles can be adopted to determine $\epsilon$, while as discussed in Tang (2018), in this case MLE is not strongly depending on the value of $\epsilon$.

the empirical return data into two groups $\mathcal{B}$ and $\mathcal{J}$, which include returns with absolute values less than or equal to $\epsilon$ and those with absolute values larger than $\epsilon$, respectively.

Here, the initial estimation of the intensity parameter, $\hat{\boldsymbol{v}}_k(dx)$, is measured as the number of jumps in period $[(k-1)h, kh]$, and for simplicity, we estimate the initial parameters $\boldsymbol{\phi}_k$ assuming that there is only one jump for a return process that belongs to group $\mathcal{J}$. Then, as discussed in Hanson and Westman (2002), the expectation and variance of the process for $t \in [(k-1)h, kh]$ are

$$\mathrm{E}(\boldsymbol{R}_t^J) = \mathrm{E}[\boldsymbol{R}_t | \boldsymbol{N}_t = 1] = \left(\boldsymbol{\mu}_k^B - \boldsymbol{\sigma}_k^{B^2}/2 - \boldsymbol{v}_k(dx)\boldsymbol{\mu}_k^J\right)\Delta + \boldsymbol{\mu}_k^J$$

and

$$\mathrm{V}(\boldsymbol{R}_t^J) = \mathrm{V}[\boldsymbol{R}_t | \boldsymbol{N}_t = 1] = \boldsymbol{\sigma}_k^{B^2}\Delta + \boldsymbol{\sigma}_k^{J^2}.$$

Hence, $\hat{\boldsymbol{\mu}}_k^J$ and $\hat{\boldsymbol{\sigma}}_k^J$ are estimated from the above equations such that

$$\begin{cases} \hat{\boldsymbol{\mu}}_k^J = \left(\hat{\mathrm{E}}\left(\boldsymbol{R}_t^J\right) - \left(\hat{\boldsymbol{\mu}}_k^B - \hat{\boldsymbol{\sigma}}_k^{B^2}/2\right)\Delta\right)\left(\mathbf{1}_A - \hat{\boldsymbol{v}}_k(dx)\Delta\right)^{-1} \\ \left(\hat{\boldsymbol{\sigma}}_k^J\right)^2 = \hat{\mathrm{V}}\left(\boldsymbol{R}_t^J\right) - \hat{\boldsymbol{\sigma}}_k^{B^2}\Delta, \end{cases}$$

where $\hat{\mathrm{E}}(\boldsymbol{R}_t^J)$ and $\hat{\mathrm{V}}(\boldsymbol{R}_t^J)$ are the sample mean and variance of the empirical returns in group $\mathcal{J}$.

When there are no jumps, the expectation and variance of the return of the process for $t \in [(k-1)h, kh]$ are

$$\mathrm{E}(\boldsymbol{R}_t^B) = \mathrm{E}[\boldsymbol{R}_t | \boldsymbol{N}_t = 0] = \left(\boldsymbol{\mu}_k^B - \boldsymbol{\sigma}_k^{B^2}/2\right)\Delta$$

and

$$\mathrm{V}(\boldsymbol{R}_t^B) = \mathrm{V}[\boldsymbol{R}_t | \boldsymbol{N}_t = 0] = \boldsymbol{\sigma}_k^{B^2}\Delta.$$

The parameters $\hat{\boldsymbol{\mu}}_k^B$ and $\hat{\boldsymbol{\sigma}}_k^B$ can be estimated from the above formulas such that

$$\begin{cases} \hat{\boldsymbol{\mu}}_k^B = \left(2\hat{\mathrm{E}}\left(\boldsymbol{R}_t^B\right) + \hat{\mathrm{V}}\left(\boldsymbol{R}_t^B\right)\Delta\right)(2\Delta)^{-1} \\ (\hat{\boldsymbol{\sigma}}_k^B)^2 = \hat{\mathrm{V}}\left(\boldsymbol{R}_t^B\right)/\Delta, \end{cases} \tag{3.27}$$

where $\hat{\mathrm{E}}(\boldsymbol{R}_t^B)$ and $\hat{\mathrm{V}}(\boldsymbol{R}_t^B)$ are the sample mean and variance of the empirical returns in group $\mathcal{B}$.

Let $\boldsymbol{R}_{\Delta t} := \ln \boldsymbol{S}_t - \ln \boldsymbol{S}_{t-\Delta}$ denote the log-return observed at $t \in [(k-1)h, kh]$, the initial estimates are then used to numerically optimize the likelihood function, given that the probability density function of returns at $\Delta t$ is:

$$\varphi_{\boldsymbol{R}_{\Delta t}}(\boldsymbol{x};\boldsymbol{\phi}_k) = \sum_{z=0}^{\infty} p_z(\boldsymbol{v}_k(dx)\Delta)\varphi_n\left(\boldsymbol{x}|\left(\boldsymbol{\mu}_k^B - {\boldsymbol{\sigma}_k^B}^2/2 - \boldsymbol{v}_k(dx)\boldsymbol{\mu}_k^J\right)\Delta + \boldsymbol{\mu}_k^J z, {\boldsymbol{\sigma}_k^B}^2\Delta + {\boldsymbol{\sigma}_k^J}^2 z^2\right),$$

where $p_z(\boldsymbol{v}_k(dx)dt) = p(d\boldsymbol{N}_t = z) = \exp(-\boldsymbol{v}_k(dx)dt)(\boldsymbol{v}_k(dx)dt)^z/z!$ for $z = 0, 1, \ldots$ and $\varphi_n$ is the normal density function (Hanson and Westman, 2002). In a multivariate economy defined in Assumption 3.4.1, returns are independent over time, so that the objective function of the MLE method is

$$L(\boldsymbol{\phi}_k) = \Pi_{j=1}^{J}\varphi_{\boldsymbol{R}_{\Delta t}}(\boldsymbol{x}_j;\boldsymbol{\phi}_k),$$

where $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_J)$ is the empirical log-return data. To estimate the five parameters, we then minimize the minus log-likelihood function:

$$-\ln L(\boldsymbol{\phi}_k) = -\sum_{j=1}^{J}\ln\varphi_{\boldsymbol{R}_{\Delta t}}(\boldsymbol{x}_j;\boldsymbol{\phi}_k).$$

Next, we numerically estimate the degree of market incompleteness at $kh$ given $\left\{\hat{\boldsymbol{\phi}}_i\right\}_{i=1}^k$ and $\left\{\hat{\boldsymbol{\phi}}_i^C\right\}_{i=1}^k$ as follows.

(i). For each asset $a = 1, 2, \ldots, A$, at time point $ih$ for $i = 1, \ldots, k$, generate 1000 replications of $dB_{n,i}^a \sim N(0,\Delta)$, $d\tilde{N}_{n,i}^a = dN_{n,i}^a - \hat{v}_i^a(dx)\Delta$ with $dN_{n,i}^a \sim Poisson(\hat{v}_i^a(dx)\Delta)$ and the observation window $\Delta = 1/252$ being the approximation for $dt$, and $\hat{J}_{n,i}^a = \exp(\hat{Q}_{n,i}^a) - 1$ with $\hat{Q}_{n,i}^a \sim N(\hat{\mu}_i^{J,a}, {\hat{\sigma}_i^{J,a}}^2)$.

(ii). For each $n = 1, 2, \ldots, 1000$ replication, calculate

$$f_n(\hat{\phi}_i^a) = \exp\left(-\hat{r}_i\Delta - g_n\left(\hat{\phi}_i^a\right)dB_{n,i}^a - \frac{1}{2}g_n^2\left(\hat{\phi}_i^a\right)\Delta\right) \times \exp\left(-d\tilde{N}_{n,i}^a - e^{-1}\hat{v}_i^a(dx)\Delta\right)$$

where

$$g_n\left(\hat{\phi}_i^a\right) = \left(\hat{\sigma}_i^{B,a}\right)^{-1}\left(\hat{\mu}_i^{B,a} - \hat{r}_i\right) - \left(\hat{\sigma}_i^{B,a}\right)^{-1}\left(\hat{J}_{n,i}^a(1-e^{-1})\hat{v}_i^a(dx)\right)$$

under incomplete markets, and

$$f_n\left(\hat{\phi}_i^{C,a}\right) = \exp\left(-\hat{r}_i\Delta - g_n\left(\hat{\phi}_i^{C,a}\right)dB_{n,i}^a - \frac{1}{2}g_n^2\left(\hat{\phi}_i^{C,a}\right)\Delta\right)$$

where $g_n\left(\hat{\phi}_i^{C,a}\right) = \left(\hat{\sigma}_i^{C,a}\right)^{-1}\left(\hat{\mu}_i^{C,a} - \hat{r}_i\right)$ with $\hat{\mu}_i^{C,a}$ and $\hat{\sigma}_i^{C,a}$ estimated following Equation (3.27) under complete markets.
(iii). Using the 1000 observations of $f_n\left(\hat{\phi}_i^a\right)$ and $f_n\left(\hat{\phi}_i^{C,a}\right)$, we find the empirical cumulative distributions $F\left(x;\hat{\phi}_i^a\right)$ and $F\left(x;\hat{\phi}_i^{C,a}\right)$ for the probability measures $P(\hat{\phi}_i^a)$ and $P(\hat{\phi}_i^{C,a})$ respectively.
(iv). Derive the distance metric for each stock $a$ at time $i$ (Frohmader and Volkmer, 2021)

$$d_3^a\left(\mathbb{m}\left(\hat{\phi}_i^a\right), \mathbb{m}\left(\hat{\phi}_i^{C,i}\right)\right) = \int_{\mathbb{R}} \left|F\left(x;\hat{\phi}_i^a\right) - F\left(x;\hat{\phi}_i^{C,a}\right)\right| dx.$$

(v). Compute the degree of market incompleteness at $kh$ for $k = 1, \ldots, K$,

$$\hat{MI}\left(\left\{\hat{\boldsymbol{\phi}}_i\right\}_{i=1}^k\right) = \frac{1}{k}\sum_{i=1}^k \frac{1}{A}\sum_{a=1}^A d_3^a\left(\mathbb{m}\left(\hat{\boldsymbol{\phi}}_i\right), \mathbb{m}\left(\hat{\boldsymbol{\phi}}_i^C\right)\right).$$

### 3.4.3 Estimation Results

Figure 3.1 displays the evolution of the degree of market incompleteness for the four stock markets. The market often sees an increase in $MI$ when there is a rising level of panic. Namely, all three developed markets experienced peaks in $MI$ during the period 2007-2009 due to the global financial crisis, in which asset prices experienced unexpected jumps due the presence of significant unhedgeable risks in the market. In a similar manner, the value of $MI$ spiked both during the mini-crash in the UK stock market in 1997 (Hua et al., 2020) as well as during the collapse of the Chinese stock market in 2015 (Han et al., 2019). Government regulation policies toward the stock market can also influence its completeness. In 1995, the sharp decline in $MI$ on the Chinese market was attributed to a policy change, which adjusted settlement dates to the next business day $(T + 1)$ instead of the same day $(T + 0)$[18] (Xu, 2000). In the Japanese market, $MI$ rose in 2000 due to deregulation policies, such as decontrolling brokerage commissions and reducing securities transaction taxes (Takaishi, 2022). We also observe that the Chinese stock market has a significantly higher degree of market incompleteness, implying that the market is susceptible to more risks that cannot be diversified away by the spanning of traded assets, which accords with the literature that emerging markets are inherently riskier (Sharkasi et al., 2006; Saranya and Prasanna, 2014).

[18] $T + 1$ came into effect on January 1, 1995, replacing $T + 0$.

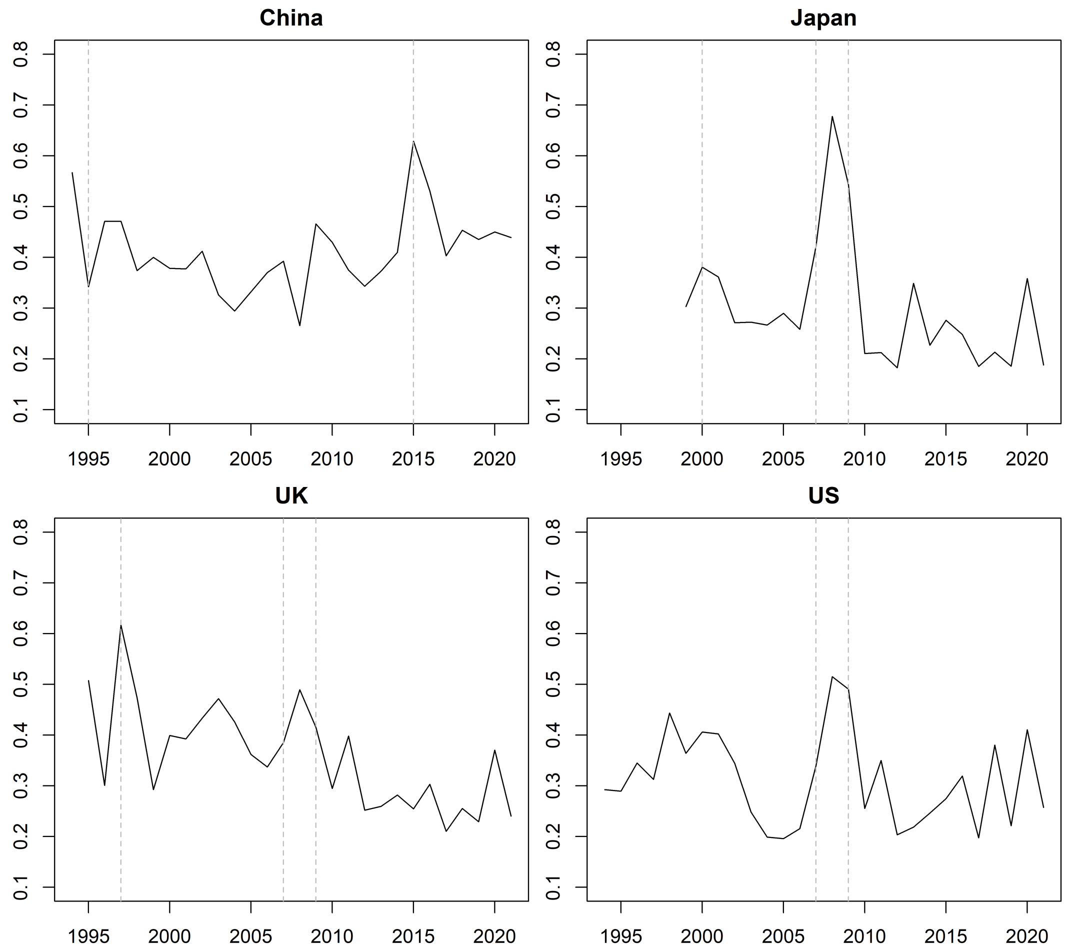


Figure 3.1: Evolution of the Degree of Market Incompleteness

## 3.5 Conclusion

This paper studies an econometric framework useful for estimating the set of SDFs in the absence of complete markets. The investigation of set properties reveals that the complete market SDF is the unique boundary point of the incomplete market SDF set, which only degenerates to its complete counterpart when the likelihood of unhedgeable risks vanish. This feature allows us to introduce a novel measure for market incompleteness, which is the distance between the probability distributions of the complete and incomplete market SDFs. We use the Wasserstein metric to construct our measure since it naturally deals with distributions with different supports.

A possible implementation of this measure is presented in which we examine the evolution of market incompleteness in the four of largest stock markets worldwide, including both emerging and developed markets. The results are consistent with our construction of incomplete markets, whereby the increase (decrease) in market incompleteness correlates to financial crises or policy changes that raise (lower) the likelihood of undiversifiable risks.

To maintain a sharp focus on our results, we have considered in detail a specific but practically realistic type of incomplete market resulting from stochastic jumps in the continuous-time setting, and applied the results in the empirical study. Nevertheless, as shown in the discrete-time setting, our framework applies more broadly, and the extension to asset prices generated by other stochastic processes is another interesting possibility worth exploring in future work. Methods of estimation and inference for more general asset-price generating processes will then refine the measurement for market incompleteness as well as the assessment of misspecification caused by imposing complete market assumptions in financial studies.

# References


Aït-Sahalia, Y. (2002): "Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach," *Econometrica*, 70, 223–262.

Aït-Sahalia, Y., J. Cacho-Diaz, and T. R. Hurd (2009): "Portfolio choice with jumps: A closed-form solution," *The Annals of Applied Probability*, 19, 556–584.

Aït-Sahalia, Y. and R. Kimmel (2007): "Maximum likelihood estimation of stochastic volatility models," *Journal of Financial Economics*, 83, 413–452.

Aït-Sahalia, Y., C. Li, and C. X. Li (2021): "Closed-form implied volatility surfaces for stochastic volatility models with jumps," *Journal of Econometrics*, 222, 364–392.

Alexander, C. and E. Lazar (2021): "The continuous limit of weak GARCH," *Econometric Reviews*, 40, 197–216.

Andersen, T. G., L. Benzoni, and J. Lund (2002): "An empirical investigation of continuous-time equity return models," *The Journal of Finance*, 57, 1239–1284.

Andersen, T. G. and T. Bollerslev (1998): "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 885–905.

Aydin, D. and E. Yilmaz (2021): "Censored nonparametric time-series analysis with autoregressive error models," *Computational Economics*, 58, 169–202.

Badescu, A., Z. Cui, and J.-P. Ortega (2017): "Non-affine GARCH option pricing models, variance-dependent kernels, and diffusion limits," *Journal of Financial Econometrics*, 15, 602–648.

Bali, T. G. and L. Peng (2006): "Is there a risk–return trade-off? Evidence from high-frequency data," *Journal of Applied Econometrics*, 21, 1169–1198.

BELLAMY, N. (2001): "Wealth optimization in an incomplete market driven by a jump-diffusion process," *Journal of Mathematical Economics*, 35, 259–287.

BERRY, T., D. GIANNAKIS, AND J. HARLIM (2015): "Nonparametric forecasting of low-dimensional dynamical systems," *Physical Review E*, 91, 032915.

BERTSIMAS, D., L. KOGAN, AND A. W. LO (2001): "Hedging derivative securities and incomplete markets: An $\epsilon$-arbitrage approach," *Operations Research*, 49, 372–397.

BIRD, R., P. PELLIZZARI, D. YEUNG, P. WOOLLEY, ET AL. (2012): "The strategic implementation of an investment process in a funds management firm," Tech. rep., Paul Woolley Centre, University of Technology Sydney and the Finance School, Waikato University.

BLUME, M. E. AND I. FRIEND (1975): "The asset structure of individual portfolios and some implications for utility functions," *The Journal of Finance*, 30, 585–603.

BOLLERSLEV, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

BONDARENKO, O. AND I. R. LONGARELA (2009): "A general framework for the derivation of asset price bounds: an application to stochastic volatility option models," *Review of Derivatives Research*, 12, 81–107.

BOSQ, D. (2012): *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, vol. 110, Springer Science & Business Media.

BOUZIANIS, G. AND L. P. HUGHSTON (2020): "Optimal Hedging in Incomplete Markets," *Applied Mathematical Finance*, 27, 265–287.

BOYLE, P., S. FENG, W. TIAN, AND T. WANG (2008): "Robust stochastic discount factors," *The Review of Financial Studies*, 21, 1077–1122.

BRANDS, S., S. J. BROWN, AND D. R. GALLAGHER (2005): "Portfolio concentration and investment manager performance," *International Review of Finance*, 5, 149–174.

BUCCHERI, G., F. CORSI, F. FLANDOLI, AND G. LIVIERI (2021): "The continuous-time limit of score-driven volatility models," *Journal of Econometrics*, 221, 655–675.

BUFFETT, W. (1994): "Letter to shareholders," *Berkshire Hathaway Annual Report*.

Buser, S. A. (1977): "Mean-variance portfolio selection with either a singular or nonsingular variance-covariance matrix," *Journal of Financial and Quantitative Analysis*, 12, 347–361.

Campbell, J. Y. (2018): *Financial Decisions and Market – A Course in Asset Pricing*, Princeton.

Campbell, R., R. Huisman, and K. Koedijk (2001): "Optimal portfolio selection in a Value-at-Risk framework," *Journal of Banking & Finance*, 25, 1789–1804.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2002): "The fine structure of asset returns: An empirical investigation," *The Journal of Business*, 75, 305–332.

Caselli, F. and J. Ventura (2000): "A representative consumer theory of distribution," *The American Economic Review*, 90, 909–926.

Cass, D. and A. Citanna (1998): "Pareto improving financial innovation in incomplete markets," *Economic Theory*, 11, 467–494.

Chen, C.-C., C.-C. Chang, E. W. Sun, and M.-T. Yu (2021): "Optimal decision of dynamic wealth allocation with life insurance for mitigating health risk under market incompleteness," *European Journal of Operational Research*.

Chen, T., J. DeJuan, and R. Tian (2018): "Distributions of GDP across versions of the Penn World Tables: A functional data analysis approach," *Economics Letters*, 170, 179–184.

Chen, T. and F. Yang (2020): "Think outside the envelop - Efficiency bound estimation through Extreme Value Theory," *Working Paper*.

Cheridito, P., U. Horst, M. Kupper, and T. A. Pirvu (2016): "Equilibrium Pricing in Incomplete Markets Under Translation Invariant Preferences," *Mathematics of Operations Research*, 41, 174–195.

Choi, N., M. Fedenia, H. Skiba, and T. Sokolyk (2017): "Portfolio concentration and performance of institutional investors worldwide," *Journal of Financial Economics*, 123, 189–208.

Christoffersen, P., R. Elkamhi, B. Feunou, and K. Jacobs (2010): "Option valuation with conditional heteroskedasticity and nonnormality," *The Review of Financial Studies*, 23, 2139–2183.

Chui, C. K. (1971): "Concerning rates of convergence of Riemann sums," *Journal of Approximation Theory*, 4, 279–287.

Claeskens, G., T. Krivobokova, and J. D. Opsomer (2009): "Asymptotic properties of penalized spline estimators," *Biometrika*, 96, 529–544.

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

Corradi, V. (2000): "Reconsidering the continuous time limit of the GARCH (1, 1) process," *Journal of Econometrics*, 96, 145–153.

Daniels, M. J. and R. E. Kass (1999): "Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models," *Journal of the American Statistical Association*, 94, 1254–1263.

Das, S. R. (2002): "The surprise element: jumps in interest rates," *Journal of Econometrics*, 106, 27–65.

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, OUP Oxford.

Dávila, E., T. Philippon, et al. (2017): "Incompleteness shocks," *Working Paper, NYU Stern.*

Dekkers, A. L., J. H. Einmahl, L. De Haan, et al. (1989): "A moment estimator for the index of an extreme-value distribution," *The Annals of Statistics*, 17, 1833–1855.

Del Barrio, E., J. A. Cuesta-Albertos, C. Matrán, and J. M. Rodríguez-Rodríguez (1999): "Tests of goodness of fit based on the L2-Wasserstein distance," *Annals of Statistics*, 1230–1239.

Dietrich, D., L. Haan, and J. Hüsler (2002): "Testing extreme value conditions," *Extremes*, 5, 71–85.

Dritschel, M. and P. Protter (1999): "Complete markets with discontinuous security price," *Finance and Stochastics*, 3, 203–214.

Duan, J.-C., P. Ritchken, and Z. Sun (2006): "Approximating GARCH-JUMP Models, Jump-Diffusion Processes, And Option Pricing," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 16, 21–52.

Durham, G. B. and A. R. Gallant (2002): "Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes," *Journal of Business & Economic Statistics*, 20, 297–338.

Ekholm, A. and B. Maury (2014): "Portfolio concentration and firm performance," *Journal of Financial and Quantitative Analysis*, 49, 903–931.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica: Journal of the Econometric Society*, 987–1007.

Evans, J. L. and S. H. Archer (1968): "Diversification and the reduction of dispersion: An empirical analysis," *The Journal of Finance*, 23, 761–767.

Fan, J. and Q. Yao (2003): *Nonlinear Time Series: Nonparametric and Parametric Methods*, vol. 20, Springer.

Fang, C.-R. and S.-Y. You (2014): "The impact of oil price shocks on the large emerging countries' stock prices: Evidence from China, India and Russia," *International Review of Economics & Finance*, 29, 330–338.

Feibel, B. J. (2003): *Investment Performance Measurement*, vol. 116, John Wiley & Sons.

Fisher, L. and J. H. Lorie (1970): "Some studies of variability of returns on investments in common stocks," *The Journal of Business*, 43, 99–134.

Floros, C. (2005): "Price linkages between the US, Japan and UK stock markets," *Financial Markets and Portfolio Management*, 19, 169–178.

Frohmader, A. and H. Volkmer (2021): "1-Wasserstein distance on the standard simplex," *Algebraic Statistics*, 12, 43–56.

Gay Jr, R. D. et al. (2008): "Effect of macroeconomic variables on stock market returns for four emerging economies: Brazil, Russia, India, and China," *International Business & Economics Research Journal (IBER)*, 7.

Geman, H. (2002): "Pure jump Lévy processes for asset price modelling," *Journal of Banking & Finance*, 26, 1297–1316.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005): "There is a risk-return trade-off after all," *Journal of Financial Economics*, 76, 509–548.

Gihman, I. I. and A. V. Skorohod (2012): *Controlled stochastic processes*, Springer Science & Business Media.

Giraitis, L., R. Leipus, and D. Surgailis (2007): "Recent advances in ARCH modelling," *Long Memory in Economics*, 3–38.

Goetzmann, W. N. and A. Kumar (2008): "Equity portfolio diversification," *Review of Finance*, 12, 433–463.

Goetzmann, W. N., L. Li, K. G. Rouwenhorst, et al. (2005): "Long-term global market correlations," *The Journal of Business*, 78, 1–38.

Goldman, E., Z. Sun, and X. Zhou (2016): "The effect of management design on the portfolio concentration and performance of mutual funds," *Financial Analysts Journal*, 72, 49–61.

Hafner, C. M., S. Laurent, and F. Violante (2017): "Weak diffusion limits of dynamic conditional correlation models," *Econometric Theory*, 33, 691–716.

Hamao, Y., R. W. Masulis, and V. Ng (1990): "Correlations in price changes and volatility across international stock markets," *The Review of Financial Studies*, 3, 281–307.

Han, C., Y. Wang, and Y. Xu (2019): "Efficiency and multifractality analysis of the Chinese stock market: Evidence from stock indices before and after the 2015 stock market crash," *Sustainability*, 11, 1699.

Hansen, L. P. and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *The Journal of Political Economy*, 99, 225–262.

Hansen, P. R. and A. Lunde (2005): "A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?" *Journal of Applied Econometrics*, 20, 873–889.

Hanson, F. B. and J. J. Westman (2002): "Stochastic analysis of jump-diffusions for financial log-return processes," in *Stochastic Theory and Control*, Springer, 169–183.

Härdle, W., H. Lütkepohl, and R. Chen (1997): "A review of nonparametric time series analysis," *International Statistical Review*, 65, 49–72.

Harrison, J. M. and D. M. Kreps (1979): "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 20, 381–408.

Harrison, J. M. and S. R. Pliska (1981): "Martingales and stochastic integrals in the theory of continuous trading," *Stochastic Processes and Their Applications*, 11, 215–260.

Heiler, S. (1999): "A survey on nonparametric time series analysis," Tech. rep., CoFE Discussion Paper.

Hua, R., P. Zhao, H. Yu, and L. Fang (2020): "Impact of US uncertainty on Chinese stock market volatility," *Emerging Markets Finance and Trade*, 56, 576–592.

Ivković, Z., C. Sialm, and S. Weisbenner (2008): "Portfolio concentration and the performance of individual investors," *Journal of Financial and Quantitative Analysis*, 43, 613–655.

Jackwerth, J. C. (2004): "Option-implied risk-neutral distributions and risk aversion," *Charlotteville: Research Foundation of AIMR.*

Jenkinson, A. F. (1955): "The frequency distribution of the annual maximum (or minimum) values of meteorological elements," *Quarterly Journal of the Royal Meteorological Society*, 81, 158–171.

Jorion, P. (2010): *Financial Risk Manager Handbook*, Wiley.

Jouzdani, N. M. and V. M. Panaretos (2021): "Functional Data Analysis with Rough Sampled Paths?" *arXiv preprint arXiv:2105.12035.*

Kacperczyk, M., C. Sialm, and L. Zheng (2005): "On the industry concentration of actively managed equity mutual funds," *The Journal of Finance*, 60, 1983–2011.

Kaido, H. and H. White (2009): "Inference on risk-neutral measures for incomplete markets," *Journal of Financial Econometrics*, 7, 199–246.

Kearney, F., F. Murphy, and M. Cummins (2015): "An analysis of implied volatility jump dynamics: Novel functional data representation in crude oil markets," *The North American Journal of Economics and Finance*, 33, 199–216.

Kent, J. T. (1983): "Information gain and a general measure of correlation," *Biometrika*, 70, 163–173.

Keynes, J. M., D. E. Moggridge, E. S. Johnson, et al. (1983): *The Collected Writings of John Maynard Keynes*, vol. XII, Macmillan London.

KLEPPE, T. S., J. YU, AND H. J. SKAUG (2014): "Maximum likelihood estimation of partially observed diffusion models," *Journal of Econometrics*, 180, 73–80.

KURISU, D. (2019): "On nonparametric inference for spatial regression models under domain expanding and infill asymptotics," *Statistics & Probability Letters*, 154, 108543.

KUSHNER, H. J. (1967): "Stochastic stability and control," Tech. rep., Brown Univ Providence RI.

KUTOYANTS, Y. A. (2012): *Identification of dynamical systems with small noise*, vol. 300, Springer Science & Business Media.

KWAK, M., T. A. PIRVU, AND H. ZHANG (2014): "A Multiperiod Equilibrium Pricing Model," *Journal of Applied Mathematics*, 2014, 1–14.

LAHIRI, S. N. (1996): "On inconsistency of estimators based on spatial data under infill asymptotics," *Sankhyā: The Indian Journal of Statistics, Series A*, 403–417.

LEDOIT, O. AND P. SANTA-CLARA (1998): "Relative pricing of options with stochastic volatility," *University of California-Los Angeles finance working paper*, 9–98.

LEDOIT, O. AND M. WOLF (2003): "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, 10, 603–621.

——— (2004): "Honey, I shrunk the sample covariance matrix," *The Journal of Portfolio Management*, 30, 110–119.

——— (2012): "Nonlinear shrinkage estimation of large-dimensional covariance matrices," *The Annals of Statistics*, 40, 1024–1060.

——— (2017): "Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks," *The Review of Financial Studies*, 30, 4349–4388.

LEITCH, G. AND J. E. TANNER (1991): "Economic forecast evaluation: profits versus the conventional error measures," *The American Economic Review*, 580–590.

LEONARD, T., J. S. HSU, ET AL. (1992): "Bayesian inference for a covariance matrix," *The Annals of Statistics*, 20, 1669–1696.

LIU, E. X. (2016): "Portfolio diversification and international corporate bonds," *Journal of Financial and Quantitative Analysis*, 51, 959–983.

LIU, K. (2017): "Effective Dimensionality Control in Quantitative Finance and Insurance," PhD thesis, University of Waterloo.

LOEB, G. M. (2007): *Battle for Investment Survival*, vol. 36, John Wiley & Sons.

LONGIN, F. (2016): *Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications*, John Wiley & Sons.

LONGIN, F. M. (2000): "From Value-at-Risk to stress testing: The extreme value approach," *Journal of Banking & Finance*, 24, 1097–1130.

MADALENO, M. AND C. PINHO (2012): "International stock market indices comovements: A new look," *International Journal of Finance & Economics*, 17, 89–102.

MALLOWS, C. L. (1972): "A note on asymptotic joint normality," *The Annals of Mathematical Statistics*, 508–515.

MARIN, J. M. AND R. RAHI (2000): "Information revelation and market incompleteness," *The Review of Economic Studies*, 67, 563–579.

MARKOWITZ, H. (1952): "Portfolio selection," *The Journal of Finance*, 7, 77–91.

MARROQUI, N., M. MORENO, ET AL. (2013): "Optimizing bounds on security prices in incomplete markets. Does stochastic volatility specification matter?" *European Journal of Operational Research*, 225, 429–442.

MATZNER-LØFBER, E., A. GANNOUN, AND J. G. DE GOOIJER (1998): "Nonparametric forecasting: a comparison of three kernel-based methods," *Communications in Statistics-Theory and Methods*, 27, 1593–1617.

MCNEIL, A. J. AND R. FREY (2000): "Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach," *Journal of Empirical Finance*, 7, 271–300.

MEDDAHI, N. (2001): "An eigenfunction approach for volatility modeling," *Cahier de Recherche*.

MEHRA, R. (2008): *Handbook of the Equity Risk Premium*, Elsevier.

MEHRA, R. AND E. C. PRESCOTT (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145–161.

Merton, R. (1972): "An analytic derivation of the efficient portfolio frontier," *Journal of Financial and Quantitative Analysis*, 7, 1851–1872.

Merton, R. C. (1976): "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, 3, 125–144.

——— (1980): "On estimating the expected return on the market: An exploratory investigation," *Journal of Financial Economics*, 8, 323–361.

Merton, R. C. and P. A. Samuelson (1992): "Continuous-time finance," Blackwell Boston.

Mnif, W. (2012): "Incomplete Market Models of Carbon Emissions Markets," *Electronic Thesis and Dissertation Repository*, 2012, 975.

Modigliani, F. and M. Leah (1997): "Risk-adjusted performance," *Journal of Portfolio Management*, 23, 45–54.

Nelson, D. B. (1990): "ARCH models as diffusion approximations," *Journal of Econometrics*, 45, 7–38.

Nguyen, X. (2011): "Wasserstein distances for discrete measures and convergence in nonparametric mixture models," *Citeseer. Forschungsbericht*.

Oyenubi, A. (2019): "Diversification measures and the optimal number of stocks in a portfolio: An information theoretic explanation," *Computational Economics*, 54, 1443–1471.

Pappas, D., K. Kiriakopoulos, and G. Kaimakamis (2010): "Optimal portfolio selection with singular covariance matrix," *International Mathematical Forum*, 5, 2305–2318.

Pena, D., G. C. Tiao, and R. S. Tsay (2011): *A Course in Time Series Analysis*, John Wiley & Sons.

Pownall, R. A. and K. G. Koedijk (1999): "Capturing downside risk in financial markets: The case of the Asian Crisis," *Journal of International Money and Finance*, 18, 853–870.

Protter, P. E. (2005): "Stochastic differential equations," in *Stochastic Integration and Differential Equations*, Springer, 249–361.

Ramsay, J. (2005): *Functional Data Analysis*, Wiley Online Library.

Rocco, M. (2014): "Extreme value theory in finance: A survey," *Journal of Economic Surveys*, 28, 82–108.

Ryabko, D. (2019): *Asymptotic Nonparametric Statistical Analysis of Stationary Time Series*, Springer.

Saranya, K. and P. K. Prasanna (2014): "Portfolio Selection and Optimization with Higher Moments: Evidence from the Indian Stock Market," *Asia-Pacific Financial Markets*, 21, 133–149.

Sharkasi, A., M. Crane, H. J. Ruskin, and J. A. Matos (2006): "The reaction of stock markets to crashes and events: A comparison study between emerging and mature markets using wavelet transforms," *Physica A: Statistical Mechanics and its Applications*, 368, 511–521.

Singleton, K. J. (2009): "Empirical dynamic asset pricing," Princeton University Press.

Skiadas, C. (2007): "Dynamic portfolio choice and risk aversion," *Handbooks in Operations Research and Management Science*, 15, 789–843.

Statman, M. (1987): "How many stocks make a diversified portfolio?" *Journal of Financial and Quantitative Analysis*, 22, 353–363.

Staum, J. (2007): "Incomplete markets," *Handbooks in Operations Research and Management Science*, 15, 511–563.

Stentoft, L. (2011): "American option pricing with discrete and continuous time models: An empirical comparison," *Journal of Empirical Finance*, 18, 880–902.

Sundaresan, S. M. (2000): "Continuous-time methods in finance: A review and an assessment," *The Journal of Finance*, 55, 1569–1622.

Svishchuk, A., D. Zhuravitskii, and A. Kalemanova (2000): "Analog of the Black-Scholes formula for option pricing under conditions of (B, S, X)-incomplete market of securities with jumps," *Ukrainian Mathematical Journal*, 52, 489–497.

Takaishi, T. (2022): "Time Evolution of Market Efficiency and Multifractality of the Japanese Stock Market," *Journal of Risk and Financial Management*, 15, 31.

Tang, F. (2018): "Merton jump-diffusion modeling of stock price data," *Linnæus University Degree project*.

Trifi, A. (2006): "Issues of aggregation over time of conditional heteroscedastic volatility models: What kind of diffusion do we recover?" *Studies in Nonlinear Dynamics & Econometrics*, 10.

Villani, C. (2009): *Optimal transport: old and new*, vol. 338, Springer.

von Mises, R. (1954): "La distribution de la plus grande de n valeurs," *American Mathematical Society, Providence, RI.*, II, 271–294.

Wang, Y. (2002): "Asymptotic nonequivalence of GARCH models and diffusions," *The Annals of Statistics*, 30, 754–783.

Willems, B. and J. Morbee (2008): "Risk management in electricity markets: hedging and market incompleteness," *TILEC Discussion Paper*.

Wu, X., H. Zhou, and S. Wang (2018): "Estimation of market prices of risks in the GARCH diffusion model," *Economic research-Ekonomska istraživanja*, 31, 15–36.

Xu, C. K. (2000): "The microstructure of the Chinese stock market," *China Economic Review*, 11, 79–97.

Yang, R. and J. O. Berger (1994): "Estimation of a covariance matrix using the reference prior," *The Annals of Statistics*, 1195–1211.

Yeung, D., P. Pellizzari, R. Bird, S. Abidin, et al. (2012): "Diversification versus Concentration... and the Winner is?" Tech. rep., Paul Woolley Centre, University of Technology Sydney and the Finance School, Waikato University.

Zhu, B., P. X.-K. Song, and J. M. Taylor (2011): "Stochastic functional data analysis: a diffusion model-based approach," *Biometrics*, 67, 1295–1304.

# APPENDICES

# Appendix A

# Appendices of Chapter 1

## A.1 The Mean-Variance Optimization

This section provides the general setup of Markowitz's mean-variance optimization. Let $\boldsymbol{E}$ be the vector of expected asset returns in the stock pool, $\mathbf{V}$ be the covariance matrix of the returns, and $\boldsymbol{w}$ be the vector of weights indicating the fraction of portfolio wealth held in each asset. Assuming that short sales are permitted, the constrained minimization problem is as follows:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \frac{1}{2}\boldsymbol{w}^T \mathbf{V} \boldsymbol{w} \\ \text{subject to} \quad & \mu = \boldsymbol{w}^T \boldsymbol{E} \text{ and } 1 = \boldsymbol{w}^T \mathbf{1}, \end{aligned}$$

where $\mu$ denotes the target expected return of the portfolio, and $\mathbf{1}$ denotes a vector of ones. The analytical solution to this problem is derived following Merton (1972), which we will not expand on here.

## A.2 The DR Method

We elaborate on the DR method first introduced by Liu (2017), which effectively reduces the number of stocks, while still preserves the variance in the market. Suppose that there are $N$ assets with asset prices $S^{(1)}, S^{(2)}, ..., S^{(N)}$ in the market. Based on the multivariate Black-Scholes model, the asset price processes $\left\{S_t^{(h)}\right\}$ for $h = 1, 2, ..., N$ solves the stochastic differential equation

$$\frac{dS_t^{(h)}}{S_t^{(h)}} = r_t dt + \sum_{l=1}^{N} \sigma_{hl} dB_t^{(l)}, \quad S_0^{(h)} = 1, \tag{A.1}$$

where $B_t^{(1)}, B_t^{(2)}, ..., B_t^{(N)}$ follow the independent standard Brownian motions, $r_t$ is the short rate of interest, and $[\sigma_{hl}]$ is the matrix capturing the correlation among the assets. Then, the solution to Equation (A.1) is

$$S_t^{(h)} = \exp\left[\left(\int_0^t r_s ds - \frac{t}{2}\sum_{l=1}^{N}\sigma_{hl}^2\right) + \sum_{l=1}^{N}\sigma_{hl}B_t^{(l)}\right], \quad h = 1, 2, ..., N. \tag{A.2}$$

Let $t_0 = 0$, $t_1 = \Delta, ..., t_m = m\Delta$ be the time steps with equal space $\Delta$, and suppose that the continuous forward rate is constant within each period. We denote $f_j$ as the annualized continuous forward rate for period $(t_{j-1}, t_j)$ such that

$$f_j = \frac{1}{\Delta}\int_{t_{j-1}}^{t_j} r_s ds, \quad j = 1, 2, ..., m.$$

Then, we have

$$\exp\left(\Delta\left(f_1 + f_2 + ... + f_j\right)\right) = \exp\left(\int_0^{t_j} r_s ds\right), \quad j = 1, 2, ..., m. \tag{A.3}$$

For $j = 1, 2, ..., m$, let $A_j^{(h)}$ be the accumulation factor of the $h^{th}$ index for the period $(t_{j-1}, t_j)$, that is,

$$A_j^{(h)} = \frac{S_{j\Delta}^{(h)}}{S_{(j-1)\Delta}^{(h)}}. \tag{A.4}$$

Combining Equation (A.2) to (A.4), we get

$$A_j^{(h)} = \exp\left[\left(f_j - \frac{1}{2}\sum_{l=1}^{N}\sigma_{hl}^2\right)\Delta + \sum_{l=1}^{N}\sigma_{hl}\sqrt{\Delta}Z_j^{(l)}\right], \tag{A.5}$$

where

$$Z_j^{(l)} = \frac{B_{j\Delta}^{(l)} - B_{(j-1)\Delta}^{(l)}}{\sqrt{\Delta}}.$$

By the property of Brownian motion, we know that $Z_1^{(l)}, Z_2^{(l)}, ..., Z_m^{(l)}$ are independent random variables with a standard normal distribution. From Equation (A.5), we derive the continuous return for the period $(t_{j-1}, t_j)$

$$R_j^{(h)} = \ln\left(A_j^{(h)}\right) = \left(f_j - \frac{1}{2}\sum_{l=1}^{N}\sigma_{hl}^2\right)\Delta + \sum_{l=1}^{N}\sigma_{hl}\sqrt{\Delta}Z_j^{(l)}.$$

The mean and covariance matrix of the returns are given by

$$\mathrm{E}\left[R_j^{(h)}\right] = \left(f_j - \frac{1}{2}\sum_{l=1}^{N}\sigma_{hl}^2\right)\Delta$$

and

$$\begin{aligned}\mathrm{Cov}\left(R_j^{(h)}, R_j^{(s)}\right) &= \mathrm{E}\left[\left(R_j^{(h)} - \mathrm{E}\left[R_j^{(h)}\right]\right)\left(R_j^{(s)} - \mathrm{E}\left[R_j^{(s)}\right]\right)\right] \\ &= \mathrm{E}\left[\left(\sum_{l=1}^{N}\sigma_{hl}\sqrt{\Delta}Z_j^{(l)}\right)\left(\sum_{l=1}^{N}\sigma_{sl}\sqrt{\Delta}Z_j^{(l)}\right)\right] \\ &= \sum_{l=1}^{N}\sigma_{hl}\sigma_{sl}\Delta, \quad h, s = 1, 2, ..., N.\end{aligned}$$

Let $\Sigma$ be the covariance matrix of the annualized continuous returns of the $N$ stocks and

$$\mathbf{A} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix}$$

be the Cholesky decomposition of $\Sigma$ such that

$$\mathbf{A}\mathbf{A}^\top = \Sigma,$$

where $\mathbf{A}^\top$ is the transpose of $\mathbf{A}$. Then, the variance contribution, also known as the explained variance (e.g., Kent, 1983), of the first $N_{DR}$ assets with the highest Sharpe Ratios can be defined as

$$\frac{\|A_1\|^2 + \dots + \|A_{N_{DR}}\|^2}{\|A_1\|^2 + \dots + \|A_N\|^2}$$

where $A_i$ is the $i^{th}$ column of $\mathbf{A}$. In this paper, the reduced dimensionality $N_{DR}$ is the minimum number of assets needed to reach the 95% explained variance, and the dimensionality reduction is achieved when $N_{DR} << N$.

## A.3 The EVT Method

In this section, we discuss the EVT method in further details. We take one stock market, say the China market, as an example and exclude all non-positive returns since our concern is the right-tail return. Let $X_1, X_2, ..., X_n$ denote the observations of returns in one group, say G1. We consider these $n$ returns as $i.i.d.$ observations from some distribution function $F$. Let $X_{1,n} \le X_{2,n} \le ... \le X_{n,n}$ be the associated order returns, so that $X_{n,n}$ denotes the maximum return in G1. Then, according to von Mises (1954) and Jenkinson (1955), if the maximum $X_{n,n}$, suitably centered and scaled, converges to a non-degenerate random variable, then there exist sequences $\{a_n\}$ $(a_n > 0)$ and $b_n$ $(b_n \in \mathbb{R})$ such that

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{X_{n,n} - b_n}{a_n} \le x\right\} = G_\gamma(x), \tag{A.6}$$

where

$$G_\gamma(x) := \exp\left(-\left(1 + \gamma x\right)^{-1/\gamma}\right)$$

for some $\gamma \in \mathbb{R}$, with $x$ such that $1 + \gamma x > 0$. That is, $F$ is in the domain of attraction of some extreme value distribution function $G_\gamma$ and $\gamma$ is the extreme-value index. By taking logarithms, Equation (A.6) can be written as

$$\lim_{q\to\infty} q\left(1 - F\left(a_q x + b_q\right)\right) = -\log G_\gamma(x) = \left(1 + \gamma x\right)^{-1/\gamma}, \quad G_\gamma(x) > 0,$$

where $q \in \mathbb{R}^+$ and $a_q$ and $b_q$ are defined by interpolation. We take $b_q = U(q)$ with

$$U(q) := \left(\frac{1}{1-F}\right)^{-1}(q) = F^{-1}\left(1 - \frac{1}{q}\right), \quad q > 1,$$

where $-1$ denotes the left-continuous inverse.
We then estimate $\gamma$, $a_q$ and $b_q$ as follows. Let, for $1 \le k < n$,

$$M_{n,k}^{(p)} := \frac{1}{k}\sum_{i=0}^{k-1}\left(\log X_{n-i,n} - \log X_{n-k,n}\right)^p, \quad p = 1, 2.$$

We use the moment estimators for $\gamma \in \mathbb{R}$ introduced by Dekkers et al. (1989):

$$\hat{\gamma} := M_{n,k}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_{n,k}^{(1)}\right)^2}{M_{n,k}^{(2)}}\right)^{-1}.$$

Specifically, we first test that $\gamma$ exists for all groups according to Dietrich et al. (2002). Next, we plot $\hat{\gamma}$ as a function of $k$, which is the number of upper order statistics used for estimation minus 1. Then, we determine the first stable region in $k$ of the estimate from the moment estimator plot. Namely, we try to identify a set of consecutive values of $k$

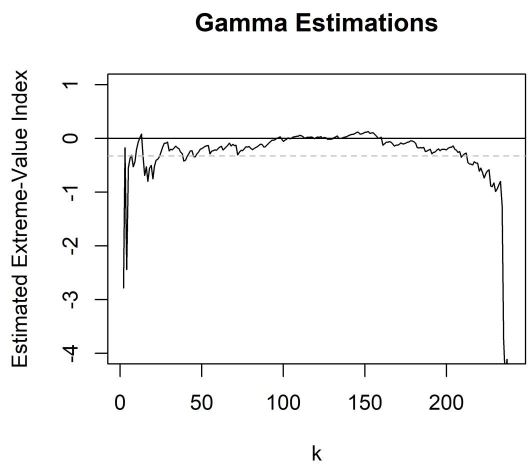


Figure A.1: Moment Estimator Versus $k$ for G1 of the China Market.

where the estimated values do not fluctuate much, so that the procedure is insensitive to the choice of $k$ in such a region. For the moment estimator in G1 for the China market, as illustrated in Figure A.1, such a stable region runs from around $k = 30$ to $k = 200$.

Next, we define the following estimators for $a_n/k$ and $b_n/k$:

$$\hat{a} := \hat{a}_{n/k} := \begin{cases} X_{n-k,n} M_{n,k}^{(1)} (1 - \hat{\gamma}) & \text{if } \hat{\gamma} < 0 \\ X_{n-k,n} M_{n,k}^{(1)} & \text{otherwise,} \end{cases}$$

and

$$\hat{b} := \hat{b}_{n/k} := X_{n-k,n}.$$

Then, our goal is to estimate the right endpoint

$$x^* := \sup \{x | F(x) < 1\}$$

of the distribution function $F$, that is, the ultimate return of G1 based on the observed returns. When estimating the endpoint, we assume that $\gamma < 0$. Next, it can be shown that Equation (A.6) is equivalent to

$$\lim_{q \to \infty} \frac{U(qx) - U(q)}{a(q)} = \frac{x^\gamma - 1}{\gamma}, \quad x > 0.$$

As $t$ gets large, we can write

$$U(qx) \approx U(q) + a(q)\frac{x^\gamma - 1}{\gamma}.$$

Because $\gamma < 0$ this yields, for large $x$ and setting $q = n/k$,

$$x^* \approx U\left(\frac{n}{k}\right) - a\left(\frac{n}{k}\right)\frac{1}{\gamma}.$$

Therefore, $x^*$ can be estimated as

$$\hat{x}^* := \hat{b} - \frac{\hat{a}}{\hat{\gamma}},$$

where $\hat{\gamma} < 0$, and $\hat{x}^* := \infty$ otherwise. The endpoint estimate of G1 for the China is shown in Figure A.2, and the selected estimate is the dotted horizontal line.

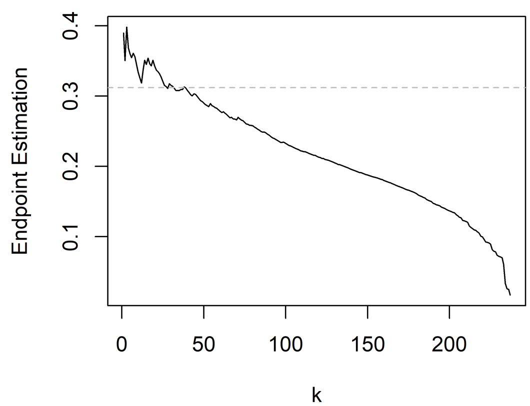


Figure A.2: Endpoint Estimators Versus $k$ for G1 of the China Market.

# Appendix B

# Appendices of Chapter 2

## B.1 Additional Notations and Lemmas.

Let $N_0(\zeta)$ and $N_{1-\Delta}(\zeta)$ denote the $\zeta$-neighborhood of $t = 0$ and $t = 1 - \Delta$ respectively, for some small $\zeta > 0$. Then we define $\mathcal{T}_{\Delta,\zeta} := [0, 1-\Delta] \cup N_0(\zeta) \cup N_{1-\Delta}(\zeta)$. Also, let $\tilde{X}_{\omega,J}$ denote the fitted function for $X_\omega$ under the sample size $J$.

**Lemma B.1.1.** *Consider a continuous sample path $X_\omega(t)$ on $[0,1]$ as such for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|t_1 - t_2| \le \delta$ implies $|X_\omega(t_1) - X_\omega(t_2)| \le \varepsilon/2$. Suppose Assumption 2.2.1 holds. Then for $B_K(t, X_\omega)$ as in Equation (2.2) and $K \ge R \ge \sup_{t\in[0,1]} |X_\omega(t)|/(\delta^2\varepsilon)$, we have the following.*

*(a) For any $t \in \mathcal{T}_{\Delta,\zeta}$ and $\omega \in \Omega$, there is $\tilde{X}_{\omega,J}(t) - B_K(t, X_\omega) \xrightarrow{p} 0$ as $J \to \infty$.*

*(b) For every $\varrho, \eta > 0$ there exists a $\xi > 0$ such that $\{\tilde{X}_{\omega,J}\}$ is asymptotically stochastically equicontinuous on $\mathcal{T}_{\Delta,\zeta}$ in that $\limsup_{J\to\infty} \mathbb{P}\{\sup_{t_1,t_2\in\mathcal{T}_{\Delta,\zeta},\ |t_1-t_2|<\zeta} |\tilde{X}_{\omega,J}(t_1) - \tilde{X}_{\omega,J}(t_2)| > \varrho\} < \eta$.*

*(c) For any $t \in \mathcal{T}_{\Delta,\zeta}$, $\omega \in \Omega$, and $r = 1, \ldots, R$, there is $\tilde{X}^{(r)}_{\omega,J_1}(t) - \tilde{X}^{(r)}_{\omega,J_2}(t) \xrightarrow{p} 0$ as $J_1, J_2 \to \infty$.*

*(d) For every $\varrho, \eta > 0$ there exists a $\zeta > 0$, such that for all $r = 1, \ldots, R-1$, there is*

$\limsup_{J\to\infty} \mathbb{P}\{\sup_{t_1,t_2\in N_0(\zeta)} |\tilde{X}^{(r)}_{\omega,J}(t_1) - \tilde{X}^{(r)}_{\omega,J}(t_2)| > \varepsilon\} < \eta$, *and*

$\limsup_{J\to\infty} \mathbb{P}\{\sup_{t_1,t_2\in N_{1-\Delta}(\zeta)} |\tilde{X}^{(r)}_{\omega,J}(t_1) - \tilde{X}^{(r)}_{\omega,J}(t_2)| > \varepsilon\} < \eta$.

**Lemma B.1.2.** *Suppose Assumption 2.2.1 and Lemma B.1.1 hold, then*

*(a)* *For* $r = 1, ..., R-1$, $\text{plim}_{J\to\infty} \lim_{t\to 0^+} \tilde{X}^{(r)}_{\omega,J}(t) = \lim_{t\to 0^+} \text{plim}_{J\to\infty} \tilde{X}^{(r)}_{\omega,J}(t)$, *and* $\text{plim}_{J\to\infty} \lim_{t\to(1-\Delta)^-} \tilde{X}^{(r)}_{\omega,J}(t) = \lim_{t\to(1-\Delta)^-} \text{plim}_{J\to\infty} \tilde{X}^{(r)}_{\omega,J}(t)$.

*(b)* $\sup_{t\in[0,1-\Delta]} |\tilde{X}_{\omega,J}(t) - B_K(t, X_\omega)| \xrightarrow{p} 0$ *as* $J \to \infty$*;*

*(c)* *For all* $r = 1, ..., R-1$, $\tilde{X}^{(r)}_{\omega,J}(0) - B^{(r)}_K(0, X_\omega) \xrightarrow{p} 0$ *and* $\tilde{X}^{(r)}_{\omega,J}(1-\Delta) - B^{(r)}_K(1-\Delta, X_\omega) \xrightarrow{p} 0$ *as* $J \to \infty$.

**Lemma B.1.3.** *Let $f$ be a function that maps a squared matrix to a real value; then for full rank squared matrices $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$, say with dimension $J$-by-$J$, there is*

$$f(\boldsymbol{Y}) = f(\boldsymbol{X}) + tr\left[\left\{\frac{\partial f(\boldsymbol{Z})}{\partial \boldsymbol{Z}}\right\}^\top (\boldsymbol{Y} - \boldsymbol{X})\right],$$

*where* $\min\{x_{ij}, y_{ij}\} < z_{ij} < \max\{x_{ij}, y_{ij}\}$ *for all elements $x_{ij}$, $y_{ij}$ and $z_{ij}$ of the matrices $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$, respectively, with $i, j = 1, ..., J$.*

## B.2 Proof of Theorem 2.2.1

First, applying integration by parts given any positive integer $R$, we have that for all $\psi \in \mathcal{F}_R$,

$$\sum_{r=1}^{R} (-1)^{r-1} \tilde{X}^{(r-1)}_{\omega,J}(1-\Delta)\psi^{(R-r)}(1-\Delta) = \sum_{r=1}^{R} (-1)^{r-1} \tilde{X}^{(r-1)}_{\omega,J}(0)\psi^{(R-r)}(0) + \int_0^{1-\Delta} \tilde{X}_{\omega,J}(t)\psi^{(R)}(t)dt - (-1)^R \int_0^{1-\Delta} \tilde{X}^{(R)}_{\omega,J}(t)\psi(t)dt,$$

$$\sum_{r=1}^{R} (-1)^{r-1} B^{(r-1)}_K(1-\Delta, X_\omega)\psi^{(R-r)}(1-\Delta) = \sum_{r=1}^{R} (-1)^{r-1} B^{(r-1)}_K(0, X_\omega)\psi^{(R-r)}(0) + \int_0^{1-\Delta} B_K(t, X_\omega)\psi^{(R)}(t)dt - (-1)^R \int_0^{1-\Delta} B^{(R)}_K(t, X_\omega)\psi(t)dt,$$

which implies that

$$
\begin{aligned}
&\left|\int_0^{1-\Delta} \tilde{X}_{\omega,J}^{(R)}(t)\psi(t)dt - \int_0^{1-\Delta} B_K^{(R)}(t,X_\omega)\psi(t)dt\right| \\
&\le \left|\sum_{r=1}^{R} (-1)^{r-1}\left\{\tilde{X}_{\omega,J}^{(r-1)}(1-\Delta) - B_K^{(r-1)}(1-\Delta,X_\omega)\right\}\psi^{(R-r)}(1-\Delta)\right| \\
&\quad + \left|\sum_{r=1}^{R} (-1)^{r-1}\left\{\tilde{X}_{\omega,J}^{(r-1)}(0) - B_K^{(r-1)}(0,X_\omega)\right\}\psi^{(R-r)}(0)\right| \\
&\quad + \left|\int_0^{1-\Delta} \tilde{X}_{\omega,J}(t)\psi^{(R)}(t)dt - \int_0^{1-\Delta} B_K(t,X_\omega)\psi^{(R)}(t)dt\right|.
\end{aligned}
$$

Lemma B.1.2(c) indicates that both $\left|\sum_{r=1}^{R} (-1)^{r-1}\left\{\tilde{X}_{\omega,J}^{(r-1)}(1-\Delta) - B_K^{(r-1)}(1-\Delta,X_\omega)\right\}\psi^{(R-r)}(1-\Delta)\right|$ and $\left|\sum_{r=1}^{R} (-1)^{r-1}\left\{\tilde{X}_{\omega,J}^{(r-1)}(0) - B_K^{(r-1)}(0,X_\omega)\right\}\psi^{(R-r)}(0)\right|$ are $o_p(1)$. Meanwhile, Lemma B.1.2(b) implies that $\left|\int_0^{1-\Delta} \tilde{X}_{\omega,J}(t)\psi^{(R)}(t)dt - \int_0^{1-\Delta} B_K(t,X_\omega)\psi^{(R)}(t)dt\right| = o_p(1)$. Hence, given any positive integer $R$,

$$
\left|\int_0^{1-\Delta} \tilde{X}_{\omega,J}^{(R)}(t)\psi(t)dt - \int_0^{1-\Delta} B_K^{(R)}(t,X_\omega)\psi(t)dt\right| = \left|\left\langle \tilde{X}_{\omega,J}^{(R)},\psi\right\rangle - \left\langle B_K^{(R)}(\cdot,X_\omega),\psi\right\rangle\right| = o_p(1),\ \ \forall\psi\in\mathcal{F}_R,
$$

and applying the properties of inner products, one can obtain

$$
\begin{aligned}
&\left\langle \hat{X}_\omega(\cdot+\Delta),\psi\right\rangle = \left\langle \sum_{r=0}^{R}\frac{1}{r!}(\Delta)^r\,\tilde{X}_{\omega,J}^{(r)},\psi\right\rangle = \sum_{r=0}^{R}\frac{1}{r!}(\Delta)^r\left\langle \tilde{X}_{\omega,J}^{(r)},\psi\right\rangle \\
&= \sum_{r=0}^{R}\frac{1}{r!}(\Delta)^r\left\langle B_K^{(r)}(\cdot,X_\omega),\psi\right\rangle + o_p(1) = \left\langle \sum_{r=0}^{R}\frac{1}{r!}(\Delta)^r B_K^{(r)}(\cdot,X_\omega),\psi\right\rangle + o_p(1) = \langle B_K(\cdot+\Delta,X_\omega),\psi\rangle + o_p(1).
\end{aligned}
$$

Then under Lemma 2.2.1, the desired results follow.

□

## B.3 Proof of Theorem 2.2.2

To improve readability, we define the followings:

$$
\mathbf{\Upsilon}_\Delta(t) := \sum_{r=0}^{R}\frac{1}{r!}\Delta^r\left[\mathbf{\Phi}^{(r)}(t)\right]^\top,\quad \mathbf{\Omega} := \frac{1}{J}\sum_{j=1}^{J}\mathbf{\Phi}(t_j)\mathbf{\Phi}^\top(t_j),\quad \mathbf{\Gamma} := \int_0^1 \mathbf{\Phi}^{(2)}(t)\left\{\mathbf{\Phi}^{(2)}(t)\right\}^\top dt.
$$

Given the order of $Q$ relative to $J$, we impose that $Q < J$ without loss of generality. Then applying Lemma B.1.3 indicates that

$$\begin{aligned}\hat{X}_{\omega}(t+\Delta) &= \boldsymbol{\Upsilon}_{\Delta}(t)\tilde{\boldsymbol{C}}_{\omega} = \boldsymbol{\Upsilon}_{\Delta}(t)\left(\boldsymbol{\Omega}+\lambda\boldsymbol{\Gamma}\right)^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)S_{t_j} = \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{\Omega}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)S_{t_j} + \mathcal{O}\left(Q\lambda\right)\\ &= \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{\Omega}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\boldsymbol{\Phi}^{\top}(t_j)\boldsymbol{C} + \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{\Omega}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\epsilon_{t_j} + \mathcal{O}\left(Q\lambda\right) + o\left(J^{-1/2}\right)\\ &= \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{C} + \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{\Omega}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\epsilon_{t_j} + \mathcal{O}\left(Q\lambda\right) + o\left(J^{-1/2}\right),\end{aligned}$$

and thus,

$$\hat{X}_{\omega}(t+\Delta) - B_K(t+\Delta, X_{\omega}) = \boldsymbol{\Upsilon}_{\Delta}(t)\boldsymbol{\Omega}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\epsilon_{t_j} + \mathcal{O}\left(Q\lambda\right) + o\left(J^{-1/2}\right).$$

Under Assumption 2.2.2(b), $\mathcal{O}\left(Q\lambda\right) = o_p\left(J^{-1/2}\right)$, and by Chui (1971),

$$\boldsymbol{\Omega} = \frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\boldsymbol{\Phi}^{\top}(t_j) = \int_0^1 \boldsymbol{\Phi}(t)\boldsymbol{\Phi}^{\top}(t)dt + o\left(J^{-1}\right).$$

Then again by Lemma B.1.3, we have

$$\hat{X}_{\omega}(t+\Delta) - B_K(t+\Delta, X_{\omega}) = \sum_{r=0}^{R}\frac{1}{r!}\Delta^r\left\{\boldsymbol{\Phi}^{(r)}(t)\right\}^{\top}\left\{\int_0^1\boldsymbol{\Phi}(s)\boldsymbol{\Phi}^{\top}(s)ds\right\}^{-1}\frac{1}{J}\sum_{j=1}^{J}\boldsymbol{\Phi}(t_j)\epsilon_{t_j} + o\left(J^{-1/2}\right),$$

and by Assumption 2.2.2(a) and the Lyapunov CLT, it follows that

$$\begin{aligned}&\left[V(t;\Delta,R,\boldsymbol{\Phi})\right]^{-1/2}\left\{\hat{X}_{\omega}(t+\Delta) - B_K(t+\Delta, X_{\omega})\right\}\\ &= \left[V(t;\Delta,R,\boldsymbol{\Phi})\right]^{-1/2}\sum_{j=1}^{J}\sum_{r=0}^{R}\frac{1}{r!}\Delta^r\left\{\boldsymbol{\Phi}^{(r)}(t)\right\}^{\top}\left\{\int_0^1\boldsymbol{\Phi}(s)\boldsymbol{\Phi}^{\top}(s)ds\right\}^{-1}\boldsymbol{\Phi}(t_j)\epsilon_{t_j} + o_p(1)\\ &= \left[V(t;\Delta,R,\boldsymbol{\Phi})\right]^{-1/2}\sum_{j=1}^{J}A_{t_j}(t;\Delta,R,\boldsymbol{\Phi})\epsilon_{t_j} + o_p(1)\\ &\xrightarrow{d}\mathcal{N}\left(0,1\right),\end{aligned}$$

where $\sigma_j^2 := \mathrm{Var}\left(\epsilon_{t_j}\right)$ for all $j$, and

$$\begin{aligned}V(t;\Delta,R,\boldsymbol{\Phi}) &:= \sum_{j=1}^{J}\sigma_j^2 A_{t_j}^2(t;\Delta,R,\boldsymbol{\Phi}),\\ A_{t_j}(t;\Delta,R,\boldsymbol{\Phi}) &:= \sum_{r=0}^{R}\frac{1}{r!}\Delta^r\left\{\boldsymbol{\Phi}^{(r)}(t)\right\}^{\top}\left\{\int_0^1\boldsymbol{\Phi}(s)\boldsymbol{\Phi}^{\top}(s)ds\right\}^{-1}\boldsymbol{\Phi}(t_j).\end{aligned}$$

Then under Lemma 2.2.1, the desired results follow with $R \to \infty$.

□

## B.4 Proof of Lemma 2.2.1

First, since $X_\omega$ is continuous on the compact set $[0,1]$, we have $c := \sup_{t\in[0,1]} |X_\omega(t)| < \infty$. Then

$$|X_\omega(t) - X_\omega(t^*)| \le c\left(\frac{t-t^*}{\delta}\right)^2 + \frac{\varepsilon}{2} \quad \forall t \in [0,1],$$

and thus,

$$\begin{aligned}
|B_K(t, X_\omega) - X_\omega(t^*)| &= |B_K(t, X_\omega - X_\omega(t^*))| && \text{(Binomial Theorem)} \\
&\le B_K\left(t, c\left(\frac{t-t^*}{\delta}\right)^2 + \frac{\varepsilon}{2}\right) \\
&= \frac{c}{\delta^2} B_K\left(t, (t-t^*)^2 + \frac{\varepsilon}{2}\right) \\
&\le \frac{c}{\delta^2}\left[t^2 + \frac{1}{K}(t-t^2) - 2tt^* + (t^*)^2\right] + \frac{\varepsilon}{2}.
\end{aligned}$$

It is implied that for $t = t^*$ specifically, we have

$$|B_K(t^*, X_\omega) - X_\omega(t^*)| \le \frac{c[t^* - (t^*)^2]}{K\delta^2} + \frac{\varepsilon}{2} \le \frac{c}{4K\delta^2} + \frac{\varepsilon}{2},$$

since $t^* - (t^*)^2 \le \frac{1}{4}$. Therefore, with $K \ge \frac{c}{2\delta^2\varepsilon}$, we have $\sup_{t\in[0,1]} |B_K(t, X_\omega) - X_\omega(t)| \le \varepsilon$.

□

## B.5 Proof of Lemma B.1.1

Lemma B.1.1(a) can be justified by the results from previous studies. For example Claeskens et al., 2009, under their Assumptions 1 to 3 that state conditions on the choice of basis functions and the distribution of the sampling points.

For part (b), with properly selected basis functions that are continuously differentiable up to a desired order, applying the mean-value theorem indicates the Lipschitz condition in that for all $t_1 < t_2 \in \mathcal{T}_{\Delta,\zeta}$ with $t_2 - t_1 < \zeta$, $\left|\tilde{X}_{\omega,J}(t_1) - \tilde{X}_{\omega,J}(t_2)\right| \le \sup_{s\in\mathcal{T}_{\Delta,\zeta}} \left|\tilde{X}^{(1)}_{\omega,J}(s)\right|(t_2 - t_1)$, where $\sup_{s\in\mathcal{T}_{\Delta,\zeta}} \left|\tilde{X}^{(1)}_{\omega,J}(s)\right| = \mathcal{O}_p(1)$. Then with the fact that $\lim_{\zeta\to 0}\sup_{|t_1-t_2|<\zeta}|t_1 - t_2| = 0$ for all $t_1, t_2 \in \mathcal{T}_{\Delta,\zeta}$, asymptotic stochastic equicontinuity of $\{\tilde{X}_{\omega,J}\}$ follows on $\mathcal{T}_{\Delta,\zeta}$.

For part (c), note that the convergence of the functional estimators is achieved through the convergence of the estimated basis coefficients, and the derivatives of these functional estimators are obtained through the derivatives of the non-stochastic basis functions; hence, the convergence of the higher order derivatives of the estimated functions can be easily justified by choosing the proper basis functions that are continuously differentiable up to a desired order.

For part (d), similarly to the justification for (b), with properly selected basis functions, one can obtain that for all $t_1 < t_2 \in N_0(\zeta)$, $\left|\tilde{X}^{(r)}_{\omega,J}(t_1) - \tilde{X}^{(r)}_{\omega,J}(t_2)\right| \le \sup_{s\in[t_1,t_2]}\left|\tilde{X}^{(r+1)}_{\omega,J}(s)\right|(t_2 - t_1)$, where $\sup_{s\in[t_1,t_2]}\left|\tilde{X}^{(r+1)}_{\omega,J}(s)\right| = \mathcal{O}_p(1)$ and $\lim_{\zeta\to 0}\sup_{t_1<t_2\in N_0(\zeta)}|t_2 - t_1| = 0$. Therefore, asymptotic stochastic equicontinuity of $\{\tilde{X}^{(r)}_{\omega,J}\}$ on $N_0(\zeta)$ for $r = 1, \dots, R-1$ follows. The same justification holds for $N_{1-\Delta}(\zeta)$.

□

## B.6 Proof of Lemma B.1.2

To verify Lemma B.1.2 (a), note that

$$\left|\lim_{t\to 0}\tilde{X}^{(r)}_{\omega,J}(t) - B^{(r)}_K(0, X_\omega)\right| \le \left|\lim_{t\to 0}\tilde{X}^{(r)}_{\omega,J}(t) - \tilde{X}^{(r)}_{\omega,J}(t)\right| + \left|\tilde{X}^{(r)}_{\omega,J}(t) - B^{(r)}_K(t, X_\omega)\right| + \left|B^{(r)}_K(t, X_\omega) - B^{(r)}_K(0, X_\omega)\right|.$$

Given any $\varrho > 0$, there exists an $\eta > 0$ such that one can find a $t$, for which there is a $\bar{J}$ that

$$\mathbb{P}\left\{\max\left\{\left|\lim_{t\to 0}\tilde{X}^{(r)}_{\omega,J}(t) - \tilde{X}^{(r)}_{\omega,J}(t)\right|, \left|\tilde{X}^{(r)}_{\omega,J}(t) - B^{(r)}_K(t, X_\omega)\right|, \left|B^{(r)}_K(t, X_\omega) - B^{(r)}_K(0, X_\omega)\right|\right\} < \varrho\right\} > 1 - \eta, \ \forall J > \bar{J},$$

implying that $\mathbb{P}\left\{\left|\lim_{t\to 0}\tilde{X}^{(r)}_{\omega,J}(t) - B^{(r)}_K(0, X_\omega)\right| < \varrho\right\} > 1 - \eta$. Therefore, with Lemma B.1.1 (c), we have that $\operatorname{plim}_{J\to\infty}\lim_{t\to 0}\tilde{X}^{(r)}_{\omega,J}(t) = \lim_{t\to 0} B^{(r)}_K(t, X_\omega) = \lim_{t\to 0}\operatorname{plim}_{J\to\infty}\tilde{X}^{(r)}_{\omega,J}(t)$.

For Lemma B.1.2 (b), applying Theorem 21.9 from Davidson (1994), with $\tilde{X}_{\omega,J}(t) - B_K(t, X_\omega) \xrightarrow{p} 0$ for each $t \in \mathcal{T}_{\Delta,\zeta}$ by Lemma B.1.1 (a), as well as asymptotic stochastic equicontinuity of $\{\tilde{X}_{\omega,J}(t)\}$ on $t \in \mathcal{T}_{\Delta,\zeta}$ by Lemma B.1.1 (b), it follows the uniform convergence in probability of $\tilde{X}_{\omega,J}$ on $\mathcal{T}_{\Delta,\zeta}$ in that $\sup_{t\in\mathcal{T}_{\Delta,\zeta}}\left|\tilde{X}_{\omega,J}(t) - B_K(t, X_\omega)\right| \xrightarrow{p} 0$. Hence, Lemma B.1.2 (b) is verified.

For Lemma B.1.2 (c), since the proofs for the two convergences follow the same idea, here we only focus on $t \to 0^+$ and omit the proof under $t \to (1-\Delta)^-$. Let $B_\zeta = N_0(\zeta) \bigcap [0, 1-\Delta]$. Then Lemma B.1.2 (c) can be proved by induction — we show that $\text{plim}_{J\to\infty} \tilde{X}^{(1)}_{\omega,J}(0) = B^{(1)}_K(0, X_\omega)$, and we justify $\text{plim}_{J\to\infty} \tilde{X}^{(r)}_{\omega,J}(0) = B^{(r)}_K(0, X_\omega)$ implies $\text{plim}_{J\to\infty} \tilde{X}^{(r+1)}_{\omega,J}(0) = B^{(r+1)}_K(0, X_\omega)$ for $r = 1, ..., R-2$.

Similarly to the verification for Lemma B.1.2 (b) based on pointwise convergence and asymptotic stochastic equicontinuity, by Lemmas B.1.1 (c) and (d) as well as Theorem 21.9 from Davidson (1994), we have $\tilde{X}^{(1)}_{\omega,J}$ converges uniformly in probability on $B_\zeta$, such that given any $\varrho > 0$ there exists an $\eta > 0$, for which one can find a $\bar{J}$ that

$$\mathbb{P}\left\{\sup_{t\in B_\zeta} \left|\tilde{X}^{(1)}_{J_1}(t) - \tilde{X}^{(1)}_{J_2}(t)\right| < \varrho\right\} > 1-\eta, \ \forall J_1, J_2 > \bar{J}. \tag{B.1}$$

Under the same $\varrho$, $\eta$, $J_1$ and $J_2$, for all $\tau \neq 0 \in B_\zeta$, applying the mean-value theorem yields

$$\left|\frac{\tilde{X}_{J_1}(\tau) - \tilde{X}_{J_2}(\tau) - \tilde{X}_{J_1}(0) + \tilde{X}_{J_2}(0)}{\tau - 0}\right| \le \sup_{t\in[0,\tau]\subset B_\zeta} \left|\tilde{X}^{(1)}_{J_1}(t) - \tilde{X}^{(1)}_{J_2}(t)\right| \le \sup_{t\in B_\zeta} \left|\tilde{X}^{(1)}_{J_1}(t) - \tilde{X}^{(1)}_{J_2}(t)\right|,$$

and with Equation (B.1), we have

$$\mathbb{P}\left\{\sup_{t\neq 0\in B_\zeta} \left|\frac{\tilde{X}_{J_1}(t) - \tilde{X}_{J_2}(t) - \tilde{X}_{J_1}(0) + \tilde{X}_{J_2}(0)}{t}\right| < \varrho\right\} > 1-\eta. \tag{B.2}$$

Define the following two functions for $t \neq 0 \in B_\zeta$:

$$g_J(t) = \frac{\tilde{X}_{\omega,J}(t) - \tilde{X}_{\omega,J}(0)}{t} \quad \text{and} \quad g(t) = \frac{B_K(t, X_\omega) - B_K(0, X_\omega)}{t};$$

then Equation (B.2) implies that $g_J$ converges uniformly in probability on $B_\zeta \backslash \{0\}$. Since $\tilde{X}_{\omega,J}$ converges uniformly to $B_K(\cdot, X_\omega)$ in probability on $B_\zeta$, it follows that

$$\operatorname*{plim}_{J\to\infty} g_J(t) = g(t), \ \forall t \neq 0 \in B_\zeta. \tag{B.3}$$

Meanwhile, given the differentiability of $\tilde{X}_{\omega,J}(t)$ and $B_K(t, X_\omega)$, we have

$$\lim_{t\to 0} g_J(t) = \tilde{X}^{(1)}_{\omega,J}(0) \text{ and } \lim_{t\to 0} g(t) = B^{(1)}_K(0, X_\omega). \tag{B.4}$$

Then, applying Lemmas B.1.2 (a) on Equations (B.3) and (B.4) indicates that $\text{plim}_{J\to\infty} \tilde{X}^{(1)}_{\omega,J}(0) = B^{(1)}_K(0, X_\omega)$.

Now suppose for a given $r$ where $r = 1, ..., R-2$, we have $\text{plim}_{J\to\infty} \tilde{X}^{(r)}_{\omega,J}(0) = B^{(r)}_K(0, X_\omega)$. Then Lemmas B.1.1 (c) and (d) as well as Theorem 21.9 from Davidson (1994) imply that $\tilde{X}^{(r+1)}_{\omega,J}$ converges uniformly in probability on $B_\zeta$, such that given any $\varrho > 0$ there exists an $\eta > 0$, for which one can find a $\bar{J}$ that $\mathbb{P}\left\{\sup_{t\in B_\zeta} \left|\tilde{X}^{(r+1)}_{J_1}(t) - \tilde{X}^{(r+1)}_{J_2}(t)\right| < \varrho\right\} > 1 - \eta$, for all $J_1, J_2 > \bar{J}$. Similarly to the previous proof, one can obtain

$$\mathbb{P}\left\{\sup_{t\neq 0\in B_\zeta}\left|\frac{\tilde{X}^{(r)}_{J_1}(t) - \tilde{X}^{(r)}_{J_2}(t) - \tilde{X}^{(r)}_{J_1}(0) + \tilde{X}^{(r)}_{J_2}(0)}{t}\right| < \varrho\right\} > 1 - \eta.$$

Note that the pointwise consistency of $\tilde{X}^{(1)}_{\omega,J}$ on $B_\zeta$ can be shown by the same means as for $\text{plim}_{J\to\infty} \tilde{X}^{(1)}_{\omega,J}(0) = B^{(1)}_K(0, X_\omega)$, switching $m$ with any point in the domain. Then re-define the following two functions for $t \neq 0 \in B_\zeta$:

$$g_J(t) = \frac{\tilde{X}^{(r)}_{\omega,J}(t) - \tilde{X}^{(r)}_{\omega,J}(0)}{t} \quad \text{and} \quad g(t) = \frac{B^{(r)}_K(t, X_\omega) - B^{(r)}_K(0, X_\omega)}{t};$$

it is implied by the uniform convergence and the pointwise consistency that $\tilde{X}^{(r)}_{\omega,J}$ converges uniformly to $B^{(r)}_K(\cdot, X_\omega)$ in probability on $B_\zeta$, it follows that

$$\operatorname*{plim}_{J\to\infty} g_J(t) = g(t), \;\; \forall t \neq 0 \in B_\zeta.$$

Meanwhile, given the differentiability of $\tilde{X}^{(r+1)}_{\omega,J}(t)$ and $B^{(r+1)}_K(t, X_\omega)$, we have

$$\lim_{t\to 0} g_J(t) = \tilde{X}^{(r+1)}_{\omega,J}(0) \text{ and } \lim_{t\to 0} g(t) = B^{(r+1)}_K(0, X_\omega).$$

Then again, applying Lemma (a) indicates that $\text{plim}_{J\to\infty} \tilde{X}^{(r+1)}_{\omega,J}(0) = B^{(r+1)}_K(0, X_\omega)$.

□

## B.7 Proof of Lemma B.1.3

First, let $\psi(q) := f(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))$ for $q \in [0,1]$. Then taking the first order derivative of $\psi(q)$ with respect to $q$ through the matrix argument of the function $f$ yields

$$\psi^{(1)}(q) = \text{tr}\left[\left\{\frac{\partial f(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))}{\partial(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))}\right\}^\top \left\{\frac{\partial(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))}{\partial q}\right\}\right] = \text{tr}\left[\left\{\frac{\partial f(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))}{\partial(\boldsymbol{X} + q(\boldsymbol{Y} - \boldsymbol{X}))}\right\}^\top (\boldsymbol{Y} - \boldsymbol{X})\right].$$

By the mean-value theorem, there exists some $q \in [0,1]$, such that $\psi(1) - \psi(0) = \psi^{(1)}(q)$, which is equivalent to

$$f(\boldsymbol{Y}) - f(\boldsymbol{X}) = \operatorname{tr}\left[\left\{\frac{\partial f(\boldsymbol{Z})}{\partial \boldsymbol{Z}}\right\}^{\top}(\boldsymbol{Y} - \boldsymbol{X})\right].$$

## B.8 FDA Results

This appendix presents functional data predictions of return and volatility, the corresponding $K$-$S$ test results with different sample sizes, and the distribution of the comparisons in relative MSFE (RMSFE). To begin with, the $K$-$S$ test results are presented in Figures B.1 and B.2 for the original pseudo-continuous sample setup with an eight-month rolling window. Next, we modify our simulation by drawing equally spaced daily points as observations from the 1000 pseudo-continuous $S_S(t)$, $v_S(t)$, $S_D(t)$ and $v_D(t)$, and keep the rolling window size of 8 months, the predictions are graphed in Figures B.3 and B.4. We then consider the case where we use the 1000 pseudo-continuous $S_S(t)$, $v_S(t)$, $S_D(t)$ and $v_D(t)$ as the sample, and reduce the rolling window to a size of 21 days (i.e., a-month-long rolling window), and the forecast results are shown in Figures B.7 and B.8. Figures B.11 and B.12 illustrate the out-of-sample forecast based on a-month-long daily data. Further, Figures B.5, B.6 (following Figures B.3 and B.4), B.9, B.10 (following Figures B.7 and B.8), B.13, and B.14 (following Figures B.11 and B.12) depict the corresponding $K$-$S$ test results, and the FDA-based method appears to correctly distinguish between the processes with different limits in out-of-sample prediction with smaller sample sizes.

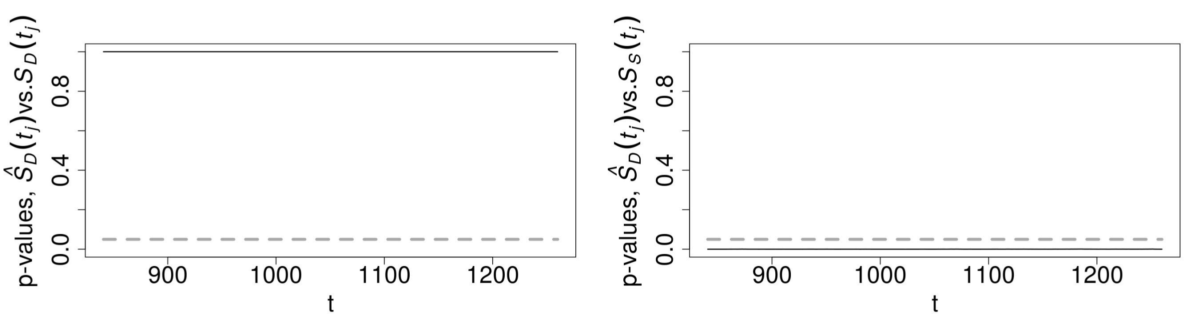


Figure B.1: Forecast vs. Underlying, $\hat{S}_D(t_j)$

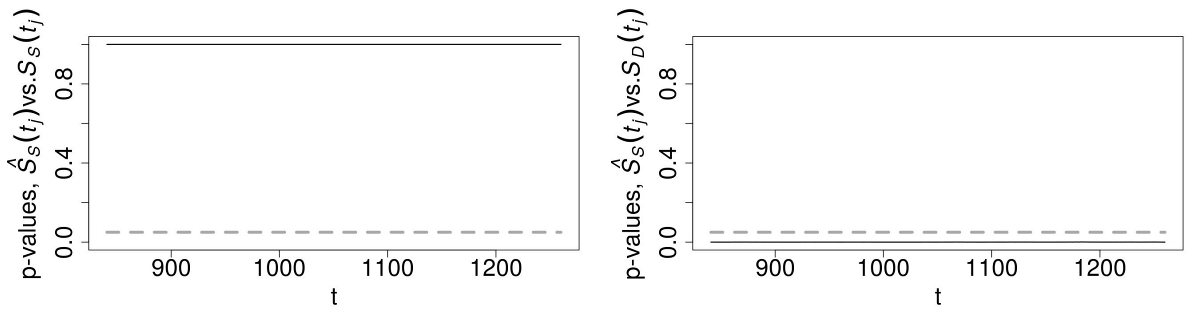


Figure B.2: Forecast vs. Underlying, $\hat{S}_S(t_j)$

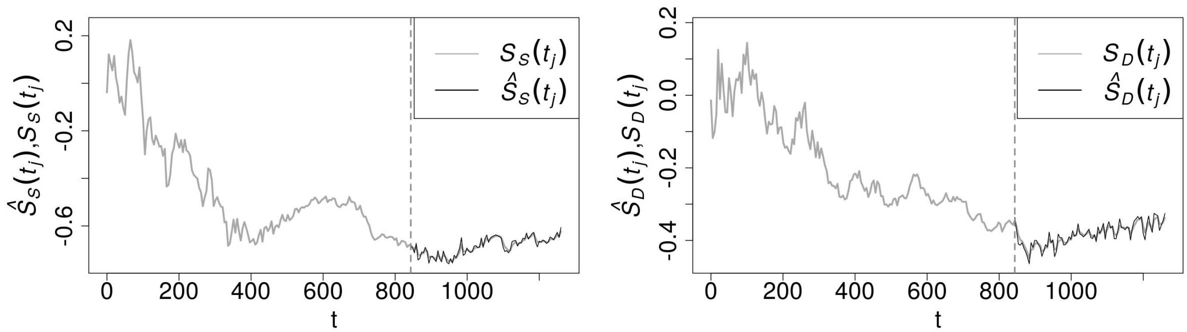


Figure B.3: Functional Data Prediction with an 8-month Rolling Window, Daily Returns

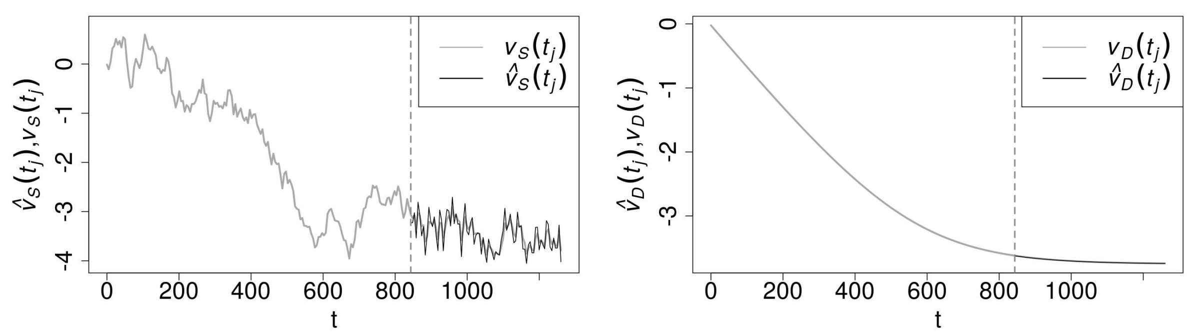


Figure B.4: Functional Data Prediction with an 8-month Rolling Window, Daily Volatility

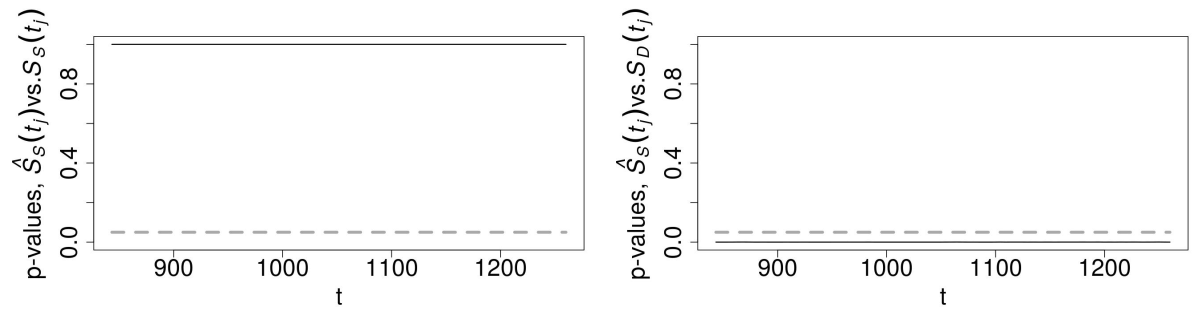


Figure B.5: Forecast vs. Underlying, Daily $\hat{S}_S(t_j)$

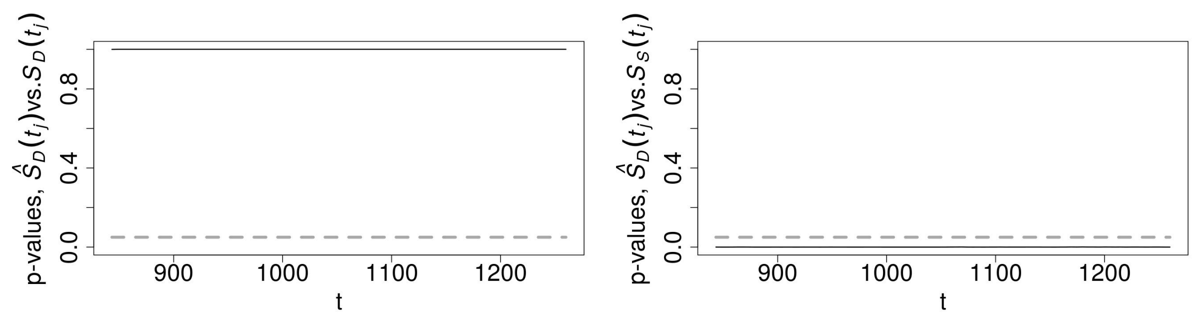


Figure B.6: Forecast vs. Underlying, Daily $\hat{S}_D(t_j)$

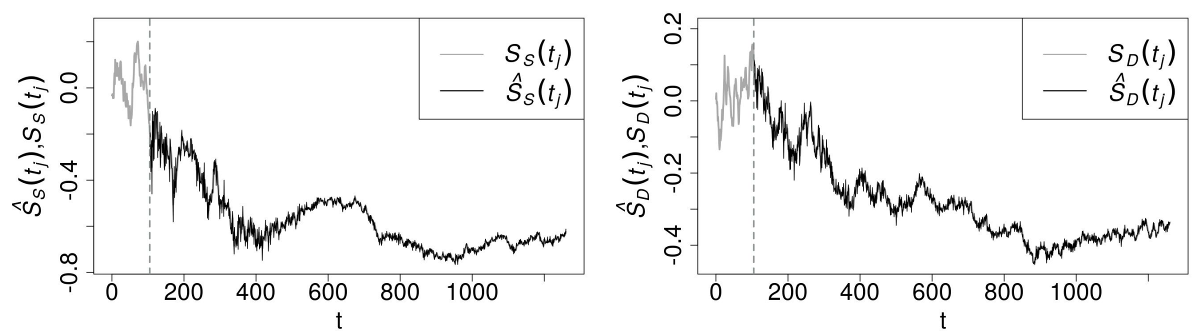


Figure B.7: Functional Data Prediction with a 1-month Rolling Window, Continuous Returns

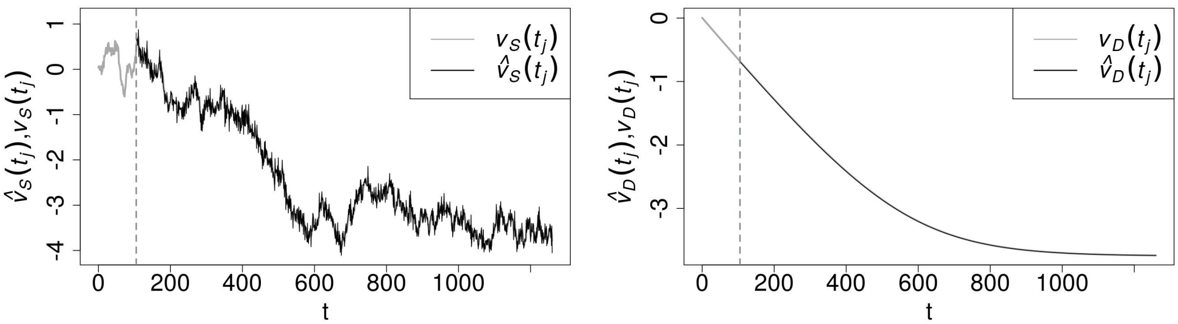


Figure B.8: Functional Data Prediction with a 1-month Rolling Window, Continuous Volatility

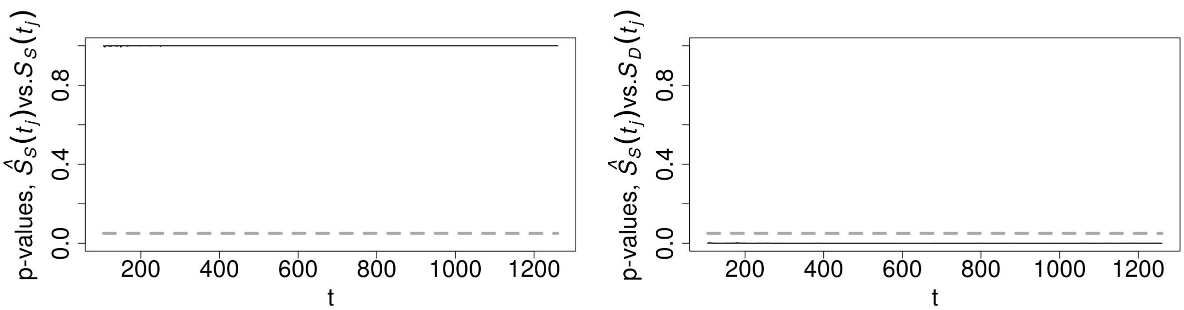


Figure B.9: Forecast vs. Underlying, Continuous $\hat{S}_S(t_j)$

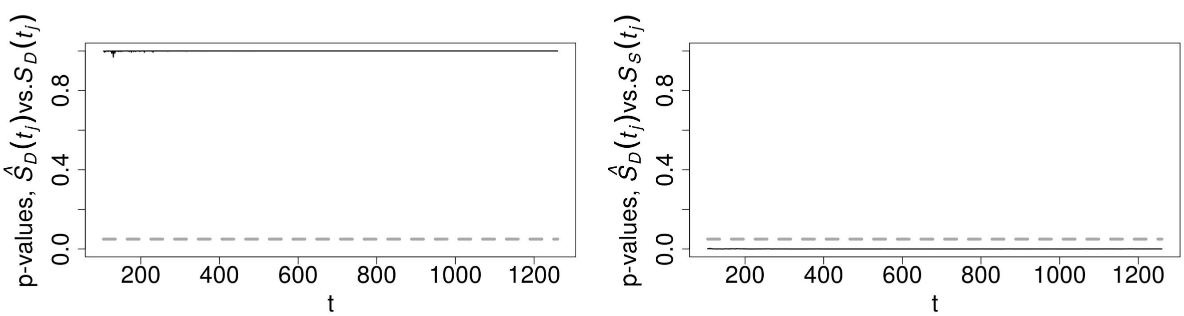


Figure B.10: Forecast vs. Underlying, Continuous $\hat{S}_D(t_j)$

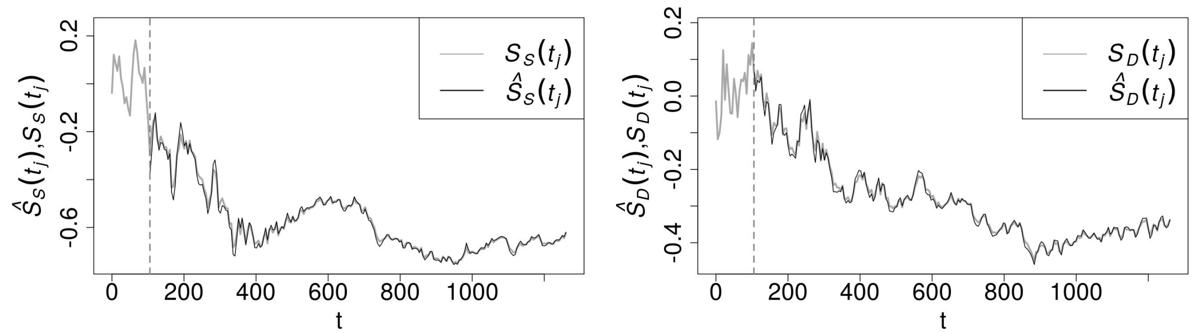


Figure B.11: Functional Data Prediction with a 1-month Rolling Window, Daily Returns

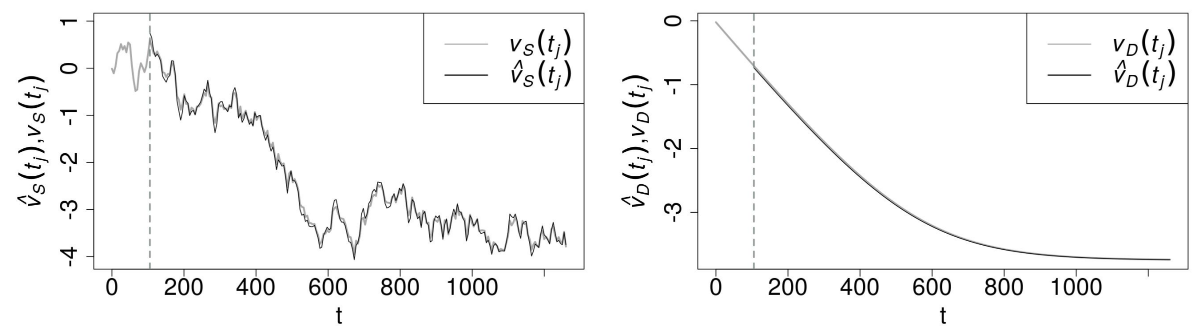


Figure B.12: Functional Data Prediction with a 1-month Rolling Window, Daily Volatility

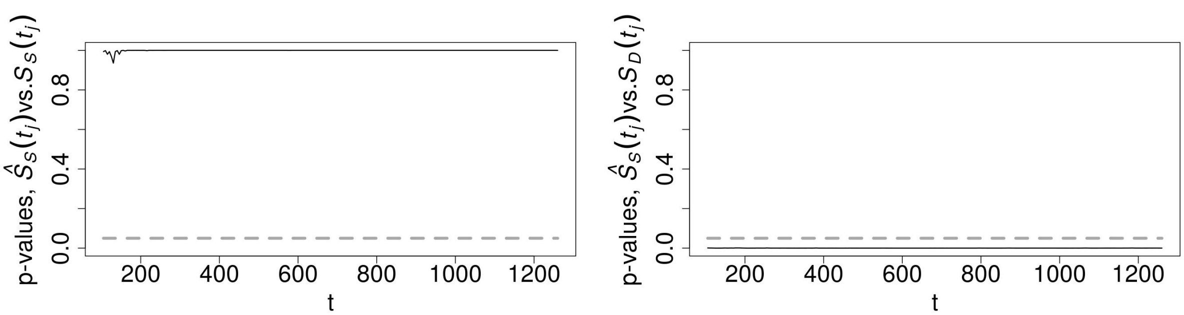


Figure B.13: Forecast vs. Underlying, Daily $\hat{S}_S(t_j)$

## B.9 MLE Results

This appendix presents detailed information on the 420 rolling window estimates, as well as the corresponding biases and rejection rates for each parameter. The horizontal dashed

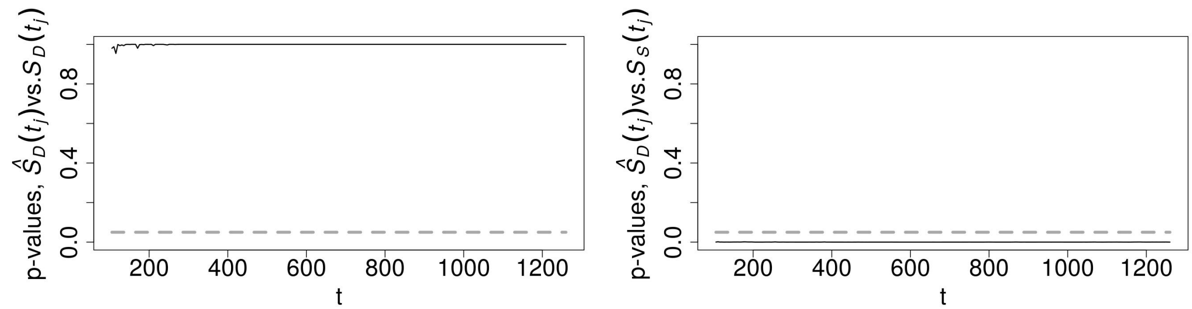


Figure B.14: Forecast vs. Underlying, Daily $\hat{S}_D(t_j)$

line indicates the true value of the parameter in the estimate plots, and the horizontal dashed line indicates the 5% significance level in the rejection rate plots. Last, Figures B.31 to B.34 present the distributions of the ratios of the MSFEs between FDA and MLE.

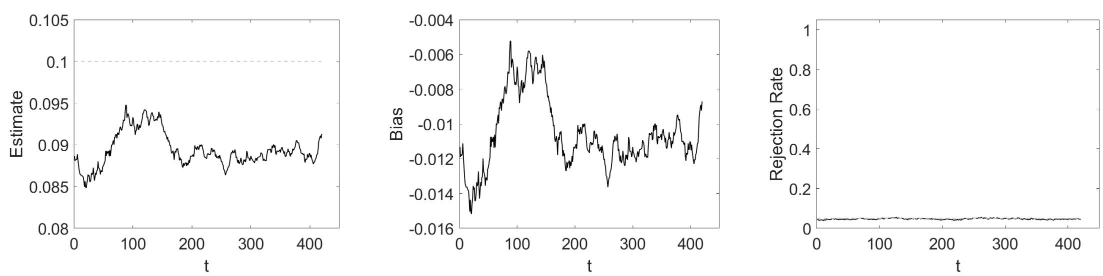


Figure B.15: Fitting DV Process Using a DV Model — $\hat{a}$

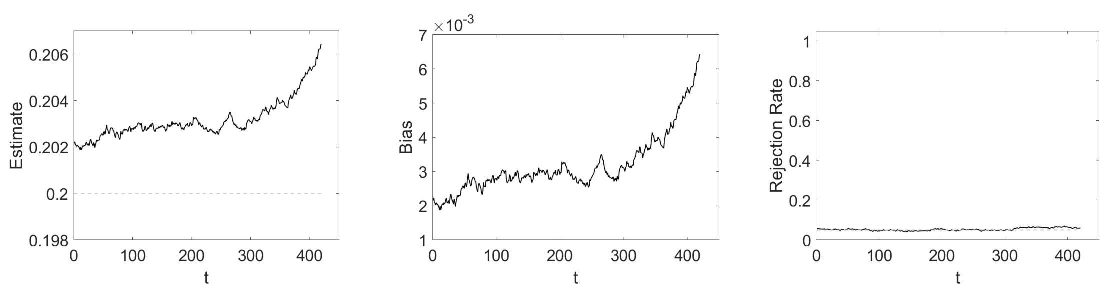


Figure B.16: Fitting DV Process Using a DV Model — $\hat{\alpha}$

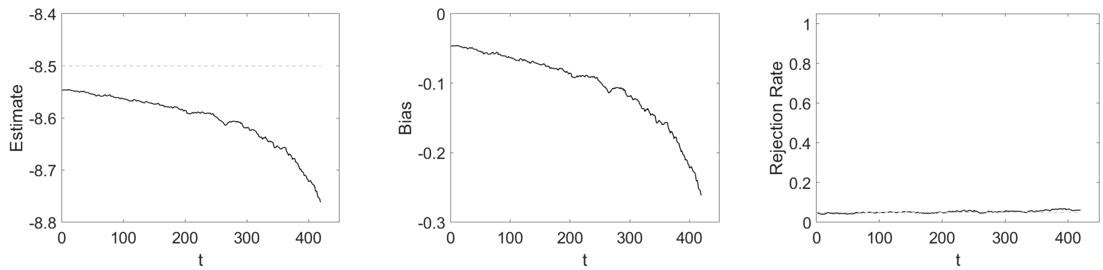


Figure B.17: Fitting DV Process Using a DV Model — $\hat{\beta}$

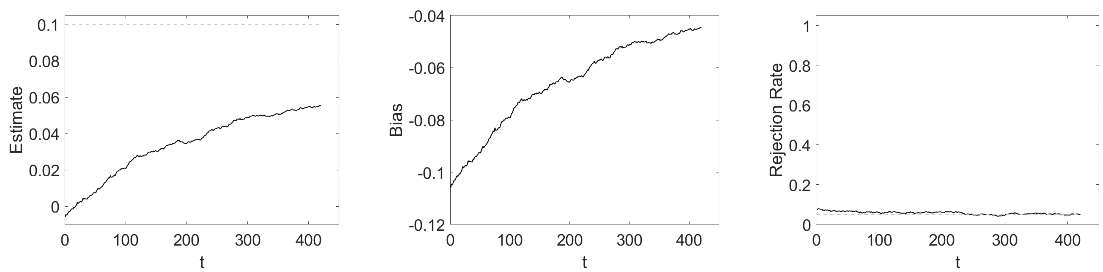


Figure B.18: Fitting SV Process Using a SV Model — $\hat{a}$

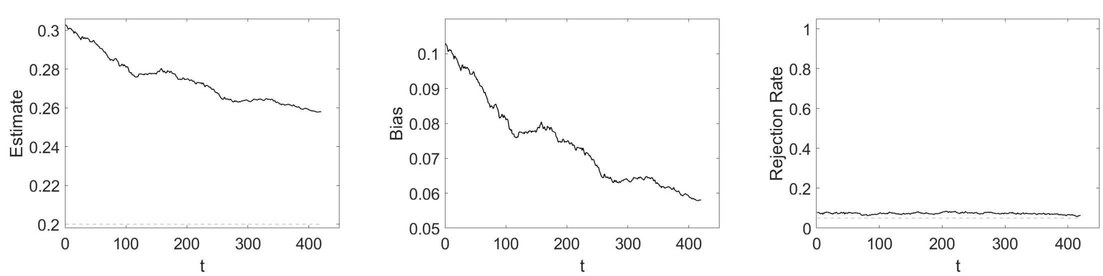


Figure B.19: Fitting SV Process Using a SV Model — $\hat{\alpha}$

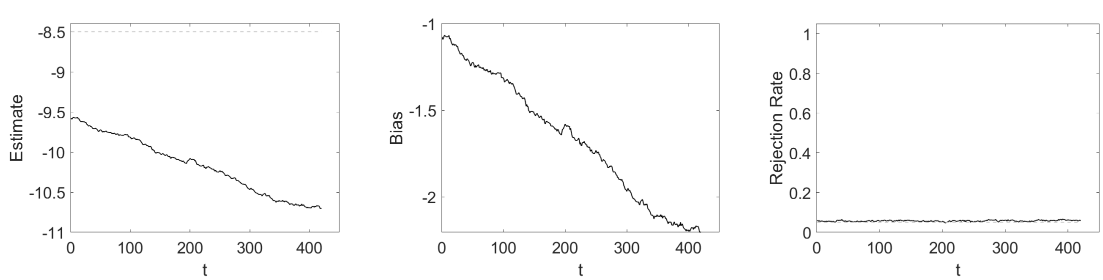


Figure B.20: Fitting SV Process Using a SV Model — $\hat{\beta}$

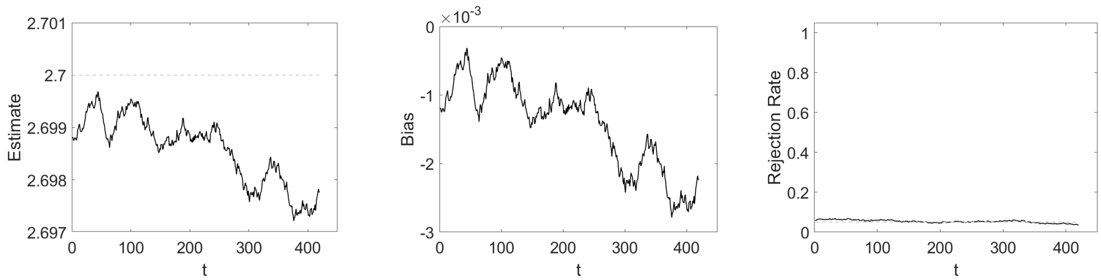


Figure B.21: Fitting SV Process Using a SV Model — $\hat{\sigma}$

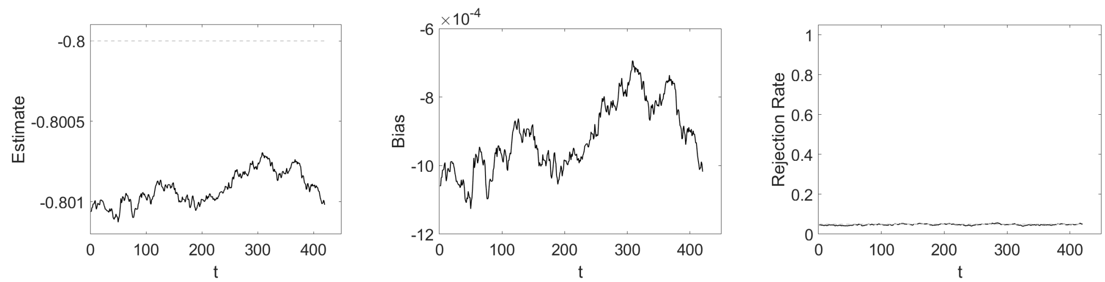


Figure B.22: Fitting SV Process Using a SV Model — $\hat{\rho}$

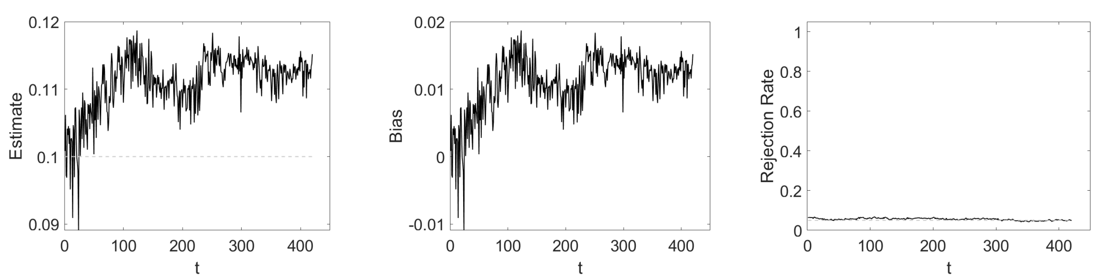


Figure B.23: Fitting SV Process Using a DV Model — $\hat{a}$

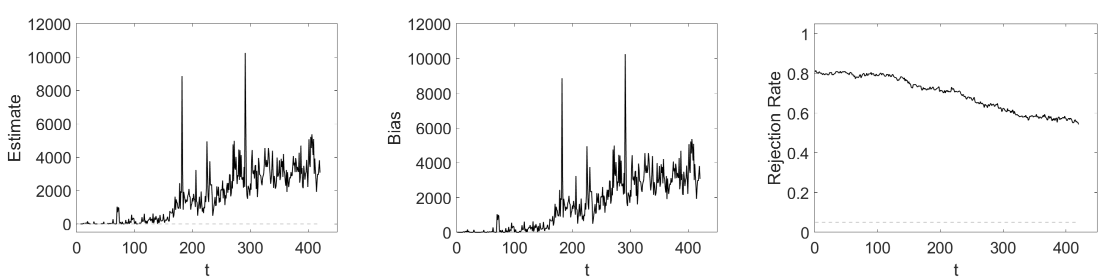


Figure B.24: Fitting SV Process Using a DV Model — $\hat{\alpha}$

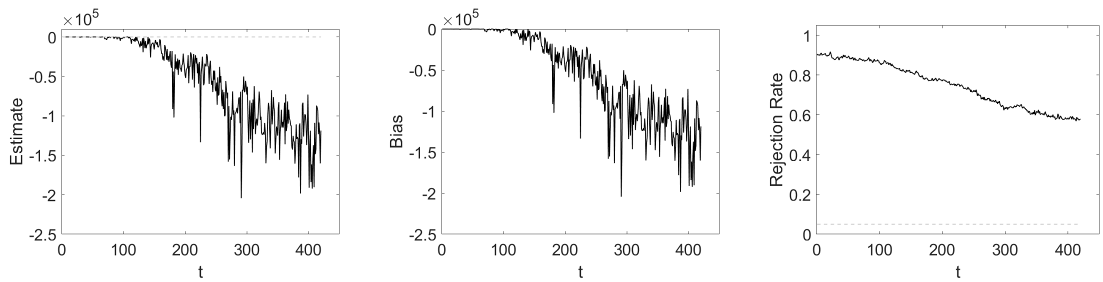


Figure B.25: Fitting SV Process Using a DV Model — $\hat{\beta}$

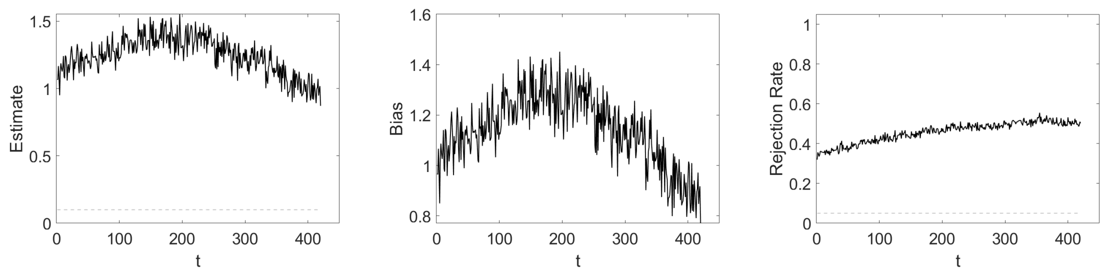


Figure B.26: Fitting DV Process Using a SV Model — $\hat{a}$

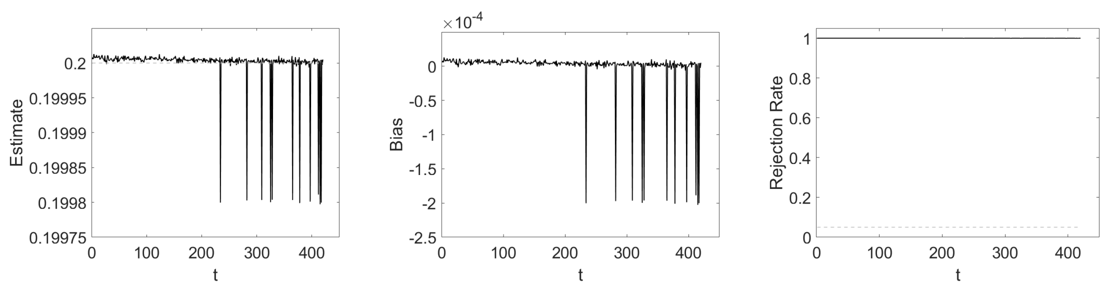


Figure B.27: Fitting DV Process Using a SV Model — $\hat{\alpha}$

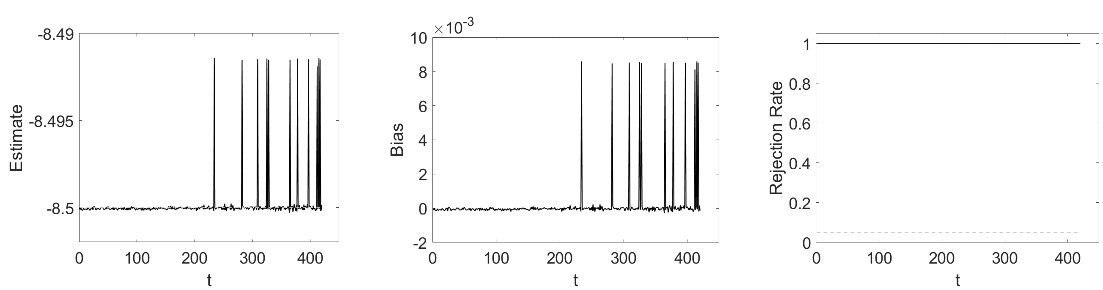


Figure B.28: Fitting DV Process Using a SV Model — $\hat{\beta}$

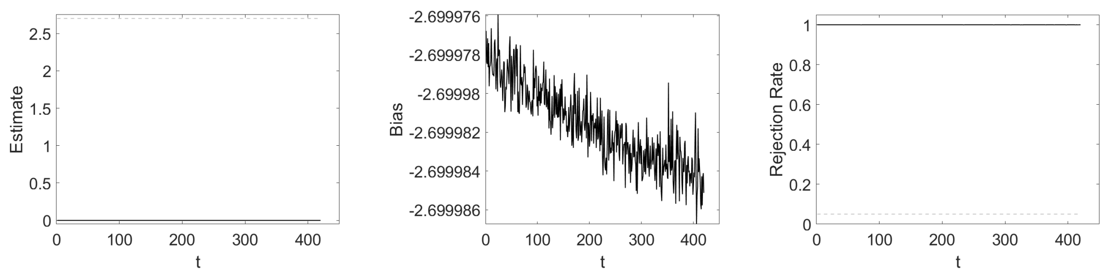


Figure B.29: Fitting DV Process Using a SV Model — $\hat{\sigma}$

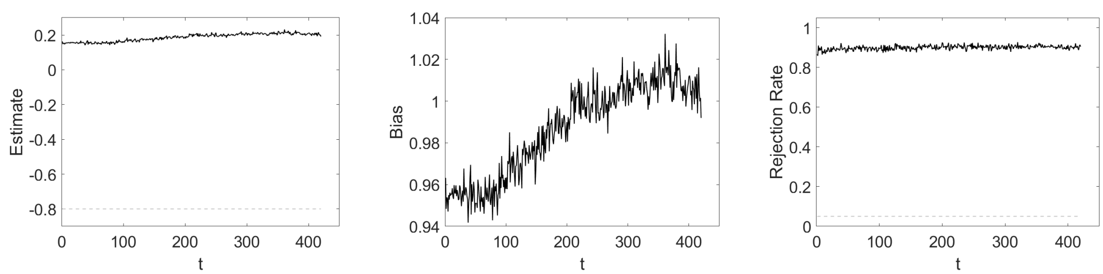


Figure B.30: Fitting DV Process Using a SV Model — $\hat{\rho}$

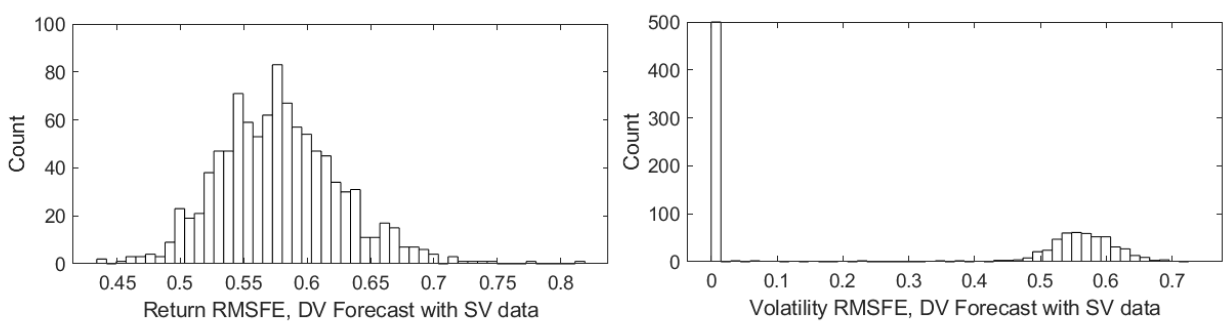


Figure B.31: FDA vs. Misspecified MLE, Relative MSFE for SV Data

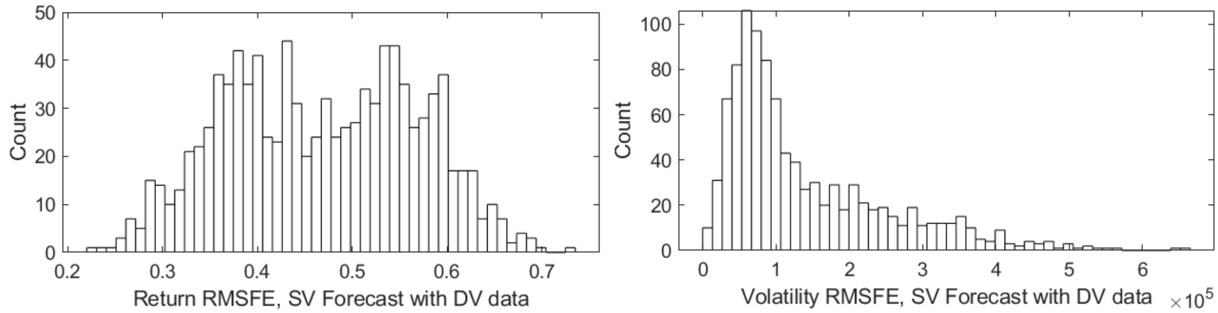


Figure B.32: FDA vs. Misspecified MLE, Relative MSFE for DV Data

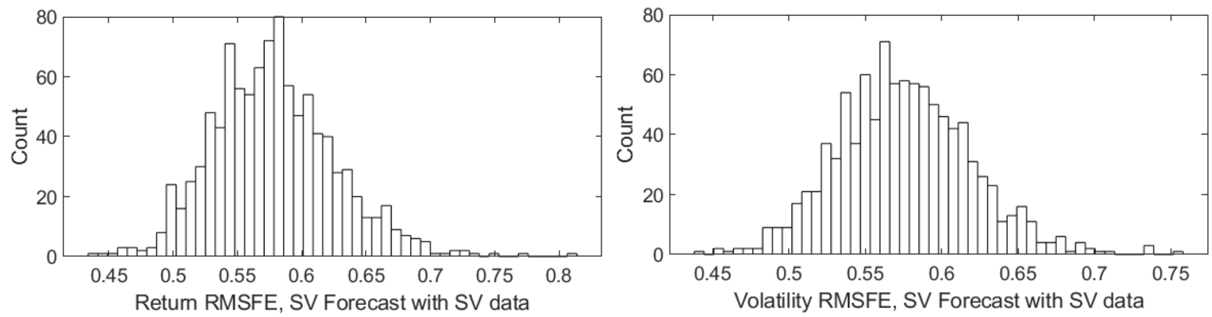


Figure B.33: FDA vs. Correctly Specified MLE, Relative MSFE for SV Data

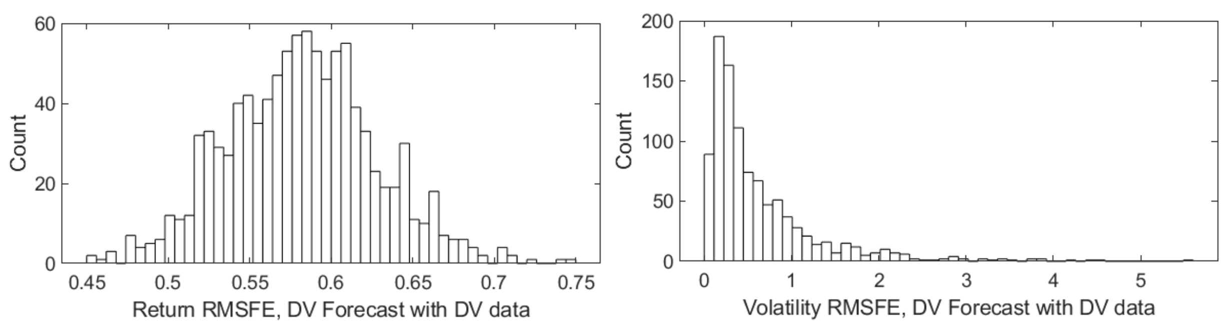


Figure B.34: FDA vs. Correctly Specified MLE, Relative MSFE for DV Data

# Appendix C

# Appendices of Chapter 3

## C.1 Proofs of Propositions and Theorems

Before proceeding to the proofs, we first recall the following definitions regarding limit and boundary points, set's convexity, openness, boundedness, and compactness properties.

**Definition C.1.1** (**Limit Point**). *Let $(\mathcal{S}, d)$ be the metric space and $\mathcal{C} \subseteq \mathcal{S}$. $x \in \mathcal{S}$ is a **limit point** of $\mathcal{C}$ if $\forall \epsilon > 0$, there is a point $y \in \mathcal{C} \setminus \{x\}$ with $d(x, y) < \epsilon$*

**Definition C.1.2** (**Boundary point**). *Let $(\mathcal{S}, d)$ be the metric space, if $\mathcal{C}$ is a subset of $\mathcal{S}$, a point $x \in \mathcal{S}$ is a **boundary point** of $\mathcal{C}$ if every neighbourhood of $x$ contains at least one point in $\mathcal{C}$ and at least one point not in $\mathcal{C}$.*

**Definition C.1.3** (**Convex set**). *Let $\mathcal{S}$ be an affine space over some ordered field. A subset $\mathcal{C}$ of $\mathcal{S}$ is **convex** if, for all $x$ and $y$ in $\mathcal{C}$, the line segment connecting $x$ and $y$ is included in $\mathcal{C}$. This means that the affine combination*

$$\rho x + (1 - \rho) y \in \mathcal{C},$$

*for all $x, y \in \mathcal{C}$, and $\rho$ in the interval $[0, 1]$.*

**Definition C.1.4** (**Open set**). *A subset $\mathcal{C}$ of a metric space $(\mathcal{S}, d)$ is **open** if every element, $x$, in $\mathcal{C}$ has a neighbourhood centred at $x$ with radius $\epsilon$ lying in the set (i.e., $\mathcal{B}(x, \epsilon) \subset \mathcal{C}$).*

**Definition C.1.5** (**Bounded set**). *A set $\mathcal{C}$ in a metric space $(\mathcal{S}, d)$ is **bounded** if it has a finite generalized diameter. In other words, there is an $R < \infty$ such that $d(x, y) \leq R$ for all $x, y \in \mathcal{C}$.*

**Definition C.1.6** (**Compact set**)**.** *For any subset $\mathcal{C}$ in a metric space $(\mathcal{S}, d)$, an* ***open cover*** *is a collection of sets $\{\boldsymbol{G}_n\}$ which are open in $(\mathcal{S}, d)$, such that $\mathcal{C} \subset \bigcup_n \{\boldsymbol{G}_n\}$. $\mathcal{C}$ is* ***compact*** *if and only if every open cover of $\mathcal{C}$ has a finite subcover.*

## C.2 Proof of Proposition 3.2.1

*Proof.* Let $\boldsymbol{x} = [x^1, x^2, x^3]^\top$ be any point in $L(\boldsymbol{\Pi})$ and $\epsilon > 0$, we prove that there is $\boldsymbol{y} = [y^1, y^2, y^3]^\top \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$ such that $d(\boldsymbol{x}, \boldsymbol{y}) < \epsilon$. Let $y^1 = x^1$, $y^2 = x^2 - \delta$, and $y^3 = x^3 + \delta$, where $\delta < \min\left\{x^2, 1 - x^3, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{s=1}^{3} y^s = 1$ and $y^s > 0$ for $s = 1, 2, 3$ imply that $\boldsymbol{y} \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$. Since $d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{2\delta^2} < \sqrt{\epsilon^2} = \epsilon$, $\boldsymbol{x} \in L(\boldsymbol{\Pi})$.

Since $\partial\boldsymbol{\Pi} \subset L(\boldsymbol{\Pi})$, every $\boldsymbol{x} \in \partial\boldsymbol{\Pi}$ is an element in $L(\boldsymbol{\Pi})$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\boldsymbol{x}, \epsilon)$ that is also an element of $\boldsymbol{\Pi}$. Now, consider $\boldsymbol{y} = [y^1, y^2, y^3]^\top \in \mathcal{B}(\boldsymbol{x}, \epsilon)$ in that $y^1 = x^1 + \delta$, $y^2 = x^2 - \delta$, and $y^3 = x^3 = 0$, where $\delta < \min\left\{1 - x^1, x^2, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{s=1}^{3} y^s = 1$, $y^1, y^2 > 0$ and $y^3 = 0$ imply that $\boldsymbol{y} \notin \boldsymbol{\Pi}$. Hence, $\boldsymbol{x} \in \partial\boldsymbol{\Pi}$. □

## C.3 Proof of Theorem 3.2.1

To prove that $d_1$ is a valid metric, we first show that, $d_2$ in Equation (3.5) is a valid metric satisfying the following conditions.

1. $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$ if and only if $M^{s_{\boldsymbol{x}}} = M^{s_{\boldsymbol{y}}}$.

   *Proof.* ($\Rightarrow$) If $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$, we must have $|v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}}| + |u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}}| = 0$ for $s_{\boldsymbol{x}} = 1, 2, \ldots, S_{\boldsymbol{x}}$ and $s_{\boldsymbol{y}} = 1, 2, \ldots, S_{\boldsymbol{y}}$. Since $M^{s_{\boldsymbol{x}}}$ and $M^{s_{\boldsymbol{y}}}$ are non-zero, we must have $v^{s_{\boldsymbol{x}}} = v^{s_{\boldsymbol{y}}}$ and $u^{s_{\boldsymbol{x}}} = u^{s_{\boldsymbol{y}}}$, and thus, $M^{s_{\boldsymbol{x}}} = M^{s_{\boldsymbol{y}}}$.

   ($\Leftarrow$) If $M^{s_{\boldsymbol{x}}} = M^{s_{\boldsymbol{y}}}$, we have $v^{s_{\boldsymbol{x}}} = v^{s_{\boldsymbol{y}}}$ and $u^{s_{\boldsymbol{x}}} = u^{s_{\boldsymbol{y}}}$, and thus, $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$. □

2. $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{x}}})$.

   *Proof.*
$$\begin{aligned} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) &= |v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}}| + |u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}}| \\ &= |v^{s_{\boldsymbol{y}}} - v^{s_{\boldsymbol{x}}}| + |u^{s_{\boldsymbol{y}}} - u^{s_{\boldsymbol{x}}}| \\ &= d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{x}}}). \end{aligned}$$
   □

3. $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{z}}}) \leq d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) + d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{z}}})$.

   *Proof.*

$$\begin{aligned}
&d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) + d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{z}}}) \\
&= |v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}}| + |v^{s_{\boldsymbol{y}}} - v^{s_{\boldsymbol{z}}}| + |u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}}| + |u^{s_{\boldsymbol{y}}} - u^{s_{\boldsymbol{z}}}| \\
&\geq |v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{y}}} + v^{s_{\boldsymbol{y}}} - v^{s_{\boldsymbol{z}}}| + |u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{y}}} + u^{s_{\boldsymbol{y}}} - u^{s_{\boldsymbol{z}}}| \\
&= |v^{s_{\boldsymbol{x}}} - v^{s_{\boldsymbol{z}}}| + |u^{s_{\boldsymbol{x}}} - u^{s_{\boldsymbol{z}}}| \\
&= d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{z}}})
\end{aligned}$$

   Hence, $d_2$ is a valid metric. □

Then, let

$$\boldsymbol{w}^* = \arg\inf_{\boldsymbol{w}} \left\{ \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) : \boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y}) \right\}$$

where $W(\boldsymbol{x}, \boldsymbol{y}) := \left\{ \boldsymbol{w} \in \mathbb{R}_+^{S_{\boldsymbol{y}} \times S_{\boldsymbol{x}}} : \boldsymbol{w}^\top \mathbf{1}_{S_{\boldsymbol{y}}} = \boldsymbol{x}, \boldsymbol{w}\boldsymbol{x} = \boldsymbol{y} \right\}$ is the set of transport plans between $\boldsymbol{x}$ and $\boldsymbol{y}$ and $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}, \mathbb{M}_{\boldsymbol{z}} \in \bar{\boldsymbol{C}}$, we prove that $d_1$ is a valid metric that satisfies the following conditions.

1. $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = 0$ if and only if $\mathbb{M}_{\boldsymbol{x}} = \mathbb{M}_{\boldsymbol{y}}$.

   *Proof.* (⇒) If $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = 0$, then we have

$$\sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0,$$

   implying that $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$ for all pairs of $(s_{\boldsymbol{x}}, s_{\boldsymbol{y}})$. Since $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \geq 0$ and $d_2$ is a valid metric, we either have $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 0$ or $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$. Since $\sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 1$, we can have one and only one $s'_{\boldsymbol{y}} \leq S_{\boldsymbol{y}}$ such that $M^{s_{\boldsymbol{x}}} = M^{s'_{\boldsymbol{y}}}$, in which case, we have $w^*_{s'_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 1$, and since $\sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} = y^{s_{\boldsymbol{y}}}$, the $s'_{\boldsymbol{y}}$ must be distinct for different $s_{\boldsymbol{x}}$. Hence, $x^{s_{\boldsymbol{x}}} = y^{s_{\boldsymbol{y}}}$ for the $s_{\boldsymbol{x}}$ such that $M^{s_{\boldsymbol{x}}} = M^{s_{\boldsymbol{y}}}$, which entails that $\mathbb{M}_{\boldsymbol{x}} = \mathbb{M}_{\boldsymbol{y}}$.

   (⇐) If $\mathbb{M}_{\boldsymbol{x}} = \mathbb{M}_{\boldsymbol{y}}$, we can have $d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) = 0$ and $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 1$ for every $s_{\boldsymbol{x}} = s_{\boldsymbol{y}}$. Since $\sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 1$ and $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \geq 0$, $w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} = 0$ for all $s_{\boldsymbol{x}} \neq s_{\boldsymbol{y}}$. Hence, $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = 0$. □

2. $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = d_1(\mathbb{M}_{\boldsymbol{y}}, \mathbb{M}_{\boldsymbol{x}})$.

*Proof.* Let $w^* \in W(\boldsymbol{x}, \boldsymbol{y})$ and $w^{*\prime} \in W(\boldsymbol{y}, \boldsymbol{x})$.

$$\begin{aligned} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \frac{w^{*\prime}_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} y^{s_{\boldsymbol{y}}}}{w^{*\prime}_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} y^{s_{\boldsymbol{y}}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^{*\prime}_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} y^{s_{\boldsymbol{y}}} \frac{w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}}}{w^{*\prime}_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} y^{s_{\boldsymbol{y}}}} d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{x}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^{*\prime}_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} y^{s_{\boldsymbol{y}}} d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{x}}}) \\ &= d_1(\mathbb{M}_{\boldsymbol{y}}, \mathbb{M}_{\boldsymbol{x}}). \end{aligned}$$

□

3. $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{z}}) \le d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) + d_1(\mathbb{M}_{\boldsymbol{y}}, \mathbb{M}_{\boldsymbol{z}})$.

*Proof.* Let $w^* \in W(\boldsymbol{x}, \boldsymbol{y})$, $w^{*\prime} \in W(\boldsymbol{y}, \boldsymbol{z})$,

$$\begin{aligned} d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{z}}) &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{z}}=1}^{S_{\boldsymbol{z}}} w^{*\prime\prime}_{s_{\boldsymbol{x}} s_{\boldsymbol{z}}} z^{s_{\boldsymbol{z}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{z}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \sum_{s_{\boldsymbol{z}}=1}^{S_{\boldsymbol{z}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} w^{*\prime}_{s_{\boldsymbol{z}} s_{\boldsymbol{y}}} y^{s_{\boldsymbol{y}}} \frac{w^{*\prime\prime}_{s_{\boldsymbol{x}} s_{\boldsymbol{z}}} z^{s_{\boldsymbol{z}}}}{w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} w^{*\prime}_{s_{\boldsymbol{z}} s_{\boldsymbol{y}}} y^{s_{\boldsymbol{y}}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{z}}}) \\ &\le \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \sum_{s_{\boldsymbol{z}}=1}^{S_{\boldsymbol{z}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \frac{w^{*\prime\prime}_{s_{\boldsymbol{x}} s_{\boldsymbol{z}}} z^{s_{\boldsymbol{z}}}}{w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) \\ &+ \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \sum_{s_{\boldsymbol{z}}=1}^{S_{\boldsymbol{z}}} w^{*\prime}_{s_{\boldsymbol{z}} s_{\boldsymbol{y}}} y^{s_{\boldsymbol{y}}} \frac{w^*_{s_{\boldsymbol{x}} s_{\boldsymbol{z}}} z^{s_{\boldsymbol{z}}}}{w^{*\prime}_{s_{\boldsymbol{z}} s_{\boldsymbol{y}}} y^{s_{\boldsymbol{y}}}} d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{z}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w^*_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) + \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \sum_{s_{\boldsymbol{z}}=1}^{S_{\boldsymbol{z}}} w^{*\prime}_{s_{\boldsymbol{z}} s_{\boldsymbol{y}}} y^{s_{\boldsymbol{y}}} d_2(M^{s_{\boldsymbol{y}}}, M^{s_{\boldsymbol{z}}}) \\ &= d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) + d_1(\mathbb{M}_{\boldsymbol{y}}, \mathbb{M}_{\boldsymbol{z}}). \end{aligned}$$

Hence, $\left(\bar{\boldsymbol{C}}, d_1\right)$ is a valid metric space. □

Now, we are ready to prove for Theorem 3.2.1.

*Proof.* Let $\mathbb{M}_{\boldsymbol{x}}$ be any point in $L(\boldsymbol{C})$, and thus, $\boldsymbol{x} = [x^1, x^2, x^3]^\top \in L(\boldsymbol{\Pi})$, and $\epsilon > 0$, we prove that there is $\mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C} \backslash \mathbb{M}_{\boldsymbol{x}}$ such that $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon$, where $\boldsymbol{y} = [y^1, y^2, y^3]^\top \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$. Let $y^1 = x^1 - \delta$, $y^2 = x^2$, and $y^3 = x^3 + \delta$. Since $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) = 0$ if and only if $\mathbb{M}_{\boldsymbol{x}} = \mathbb{M}_{\boldsymbol{y}}$ and $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) \geq 0$, thus, we can choose $\delta$ satisfying the following conditions:

$$d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon \text{ and } \delta < \min\left\{x^1, 1 - x^3\right\},$$

so that we have $\sum_{s=1}^{3} y^s = 1$ and $y^s > 0$ for $s = 1, 2, 3$, implying that $\boldsymbol{y} \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$. Therefore, $\mathbb{M}_{\boldsymbol{y}} \in \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ such that $\mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C} \backslash \mathbb{M}_{\boldsymbol{x}}$, and thus, $\mathbb{M}_{\boldsymbol{x}} \in L(\boldsymbol{C})$.

Since $\partial \boldsymbol{C} \subset L(\boldsymbol{C})$, every $\mathbb{M}_{\boldsymbol{x}} \in \partial \boldsymbol{C}$ is an element in $L(\boldsymbol{C})$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ that is also an element of $\boldsymbol{C}$. Now, consider $\mathbb{M}_{\boldsymbol{y}} \in \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ in that $y^1 = x^1 + \delta$, $y^2 = x^2 - \delta$, and $y^3 = x^3 = 0$, where $\delta < \min\{1 - x^1, x^2\}$ and satisfies the following condition $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon$. Then, $y^s > 0$ for $s = 1, 2$, and $y^3 = 0$, implying that $\boldsymbol{y} \notin \boldsymbol{\Pi}$, and thus, $\mathbb{M}_{\boldsymbol{y}} \notin \boldsymbol{C}$. Hence, $\mathbb{M}_{\boldsymbol{x}} \in \partial \boldsymbol{C}$. □

## C.4 Proof of Lemma 3.2.1

*Proof.* Suppose that $(M_n^1, M_n^2) \to (M^1, M^2) \in \partial \boldsymbol{C}$ with $\pi_n^1 + \pi_n^2 = 1$ and $1/\pi_n^1, 1/\pi_n^2 > 0$. If it were $1/\pi^1 = 0$, then $\pi_n^1 \to \infty$, but $\pi_n^1 \leq \pi_n^1 + \pi_n^2 = 1$, so that's impossible. Similarly, we cannot have $1/\pi_n^2 = 0$. Then, $\lim_{n\to\infty} \pi_n^1 + \pi_n^2 = \pi^1 + \pi^2 = 1$ so that $(M_n^1, M_n^2) \in \partial \boldsymbol{C}$, and $\partial \boldsymbol{C}$ is closed. Since $\partial \boldsymbol{C}$ is a non-empty subspace of $\bar{\boldsymbol{C}}$, taking $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, there exists a closed ball $\mathcal{B} = \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ such that $\mathcal{B} \bigcap \partial \boldsymbol{C}$ is a non-empty compact set. So the function $\mathbb{M}_{\boldsymbol{y}} \mapsto d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$ defined on $\mathcal{B} \bigcap \partial \boldsymbol{C}$ must achieve a minimum. That is, there is some $\mathbb{M}_{\boldsymbol{y}} = \mathbb{M}_{\boldsymbol{y}}^* \in \mathcal{B} \bigcap \partial \boldsymbol{C}$, which minimizes $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$. Further, for $\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C} \backslash \mathcal{B}$, we have $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) > \epsilon \geq d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}^*)$, so it minimizes the distance on the whole of $\partial \boldsymbol{C}$. Moreover, since $\mathbb{M}_{\boldsymbol{y}}^* \notin \boldsymbol{C}$, $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}^*) > 0$. Hence, for every $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, there exists $\mathbb{M}_{\boldsymbol{y}}^*$ such that

$$\mathbb{M}_{\boldsymbol{y}}^* = \underset{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}}{\arg\min} \{d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})\}.$$

□

## C.5 Proof of Theorem 3.2.2

*Proof.* Let $\rho \in [0, 1]$ and $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C}$. The affine combination of $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}$ is

$$\begin{aligned}\rho\,\mathbb{M}_{\boldsymbol{x}} + (1-\rho)\,\mathbb{M}_{\boldsymbol{y}} &= \{\boldsymbol{M} : \rho \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{r}\boldsymbol{M}] + (1-\rho)\mathrm{E}_{\boldsymbol{y}}[\boldsymbol{r}\boldsymbol{M}] = \rho \boldsymbol{1}_2 + (1-\rho)\boldsymbol{1}_2\} \\ &= \{\boldsymbol{M} : \rho \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{r}\boldsymbol{M}] + (1-\rho)\mathrm{E}_{\boldsymbol{y}}[\boldsymbol{r}\boldsymbol{M}] = \boldsymbol{1}_2\} \in \boldsymbol{C}.\end{aligned}$$

Hence, $\boldsymbol{C}$ is *convex.*

Let $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$ with $\boldsymbol{x} = [x^1, x^2, x^3]^\top \in \boldsymbol{\Pi}$. There is $\eta > 0$ such that

$$\eta = \min_{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}} \{d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})\},$$

where $\boldsymbol{y} \in \partial \boldsymbol{\Pi}$. Then, since $\boldsymbol{C} = \bar{\boldsymbol{C}} \backslash \partial \boldsymbol{C}$, by choosing $\epsilon < \eta$, we have $\mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon) \subset \boldsymbol{C}$. Hence, $\boldsymbol{C}$ is *open* in $\left(\bar{\boldsymbol{C}}, d_1\right)$.

Next, we prove that $\boldsymbol{C}$ is bounded in $\left(\bar{\boldsymbol{C}}, d_1\right)$. First, let $\boldsymbol{\kappa_v}$ and $\boldsymbol{\kappa_u}$ denote the coefficient vectors of $\boldsymbol{v}(\boldsymbol{x})$ and $\boldsymbol{u}(\boldsymbol{x})$, respectively, for $\boldsymbol{x} \in \bar{\boldsymbol{\Pi}}$. That is

$$\begin{cases} \boldsymbol{\kappa_v} = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \\ 0 \end{bmatrix} \text{ and } \boldsymbol{\kappa_u} = \begin{bmatrix} \frac{r^2-r^3}{r^1-r^2} \\ -\frac{r^1-r^3}{r^1-r^2} \\ 1 \end{bmatrix}, & \boldsymbol{x} \in \boldsymbol{\Pi}; \\ \boldsymbol{\kappa_v} = \begin{bmatrix} \frac{1-r^2}{r^1-r^2} \\ \frac{r^1-1}{r^1-r^2} \end{bmatrix} \text{ and } \boldsymbol{\kappa_u} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & \boldsymbol{x} \in \partial \boldsymbol{\Pi}. \end{cases}$$

Then, given that $\boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y})$, we have

$$\begin{aligned} \mathcal{D} &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \left( |\kappa_{\boldsymbol{v}}^{s_{\boldsymbol{x}}}/x^{s_{\boldsymbol{x}}} - \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}}/y^{s_{\boldsymbol{y}}}| + |\kappa_{\boldsymbol{u}}^{s_{\boldsymbol{x}}}/x^{s_{\boldsymbol{x}}} - \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{y}}}/y^{s_{\boldsymbol{y}}}| \right) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \left( |w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{x}}} - w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}}| + |w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{x}}} - w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}}| \right). \end{aligned}$$

Since $\sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} = y^{s_{\boldsymbol{y}}}$, for every $s_{\boldsymbol{x}} = 1, 2, \ldots, S_{\boldsymbol{x}}, s_{\boldsymbol{y}} = 1, 2, \ldots, S_{\boldsymbol{y}}$, $w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \in [0, y^{s_{\boldsymbol{y}}}]$, and thus, $w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}} \in [0, \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}}]$. Therefore, $\mathcal{D}$ is bounded, implying that there is $0 < R < \infty$ such that $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) \leq R$ for all $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C}$. Hence, $\boldsymbol{C}$ is *bounded.*

Lastly, to show that $\boldsymbol{C}$ is not compact, we just need one example of an open cover that has no finite open subcovers. Let $\{\boldsymbol{G}_n\} = \{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \boldsymbol{\Pi}_n, n \in \mathbb{N}\}$, where

$$\{\boldsymbol{\Pi}_n\} = \left\{ \left[ \pi^1, \pi^2, 1 - \sum_{i=1}^{2} \pi^s \right]^\top \in \mathbb{R}^3_{++} : \sum_{i=1}^{2} \pi^s \in \left( \frac{1}{n}, \frac{n-1}{n} \right), n \in \mathbb{N} \right\}.$$

Notice that, if this gives us an invalid segment such as $(1, 0)$, we treat it as an empty element. Here, for any $b = \sum_{i=1}^{2} \pi^s \in (0, 1)$, the Archimedean Property provides an $n \in \mathbb{N}$ such that $n > \max\left\{\frac{1}{b}, \frac{1}{1-b}\right\}$. Then,

$$\begin{aligned} &nb > 1 \text{ and } n - nb > 1 \\ \Rightarrow\ &1 < nb < n - 1 \\ \Rightarrow\ &b \in \left(\frac{1}{n}, \frac{n}{n-1}\right). \end{aligned}$$

Thus, every element of $\mathbf{\Pi}$ is in $\{\mathbf{\Pi}_n\}$ for some $n \in \mathbb{N}$, and therefore, every element of $\boldsymbol{C}$ is in $\{\boldsymbol{G}_n\}$ for some $n \in \mathbb{N}$, suggesting that $\boldsymbol{C} \subset \bigcup_{n=1}^{\infty} \{\boldsymbol{G}_n\}$. Moreover, since for any $n \in \mathbb{N}$, $\boldsymbol{G}_n$ has a neighbourhood centred at $\boldsymbol{G}_n$ with radius $\epsilon > 0$ lying in the set, $\{\boldsymbol{G}_n\}$ is an open cover of $\boldsymbol{C}$. Let $k, l \in \mathbb{N}$ such that $k > l > 2$, we have

$$\frac{1}{k} < \frac{1}{l} < 1 - \frac{1}{l} < 1 - \frac{1}{k} \Rightarrow \{\mathbf{\Pi}_l\} \subset \{\mathbf{\Pi}_k\} \Rightarrow \{\boldsymbol{G}_l\} \subset \{\boldsymbol{G}_k\}.$$

Therefore, for any finite $m \in \mathbb{N}$, $\bigcup_{n=1}^{m} \{\boldsymbol{G}_n\} = \{\boldsymbol{G}_m\} = \{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_m, m \in \mathbb{N}\}$. However, for any $m \in \mathbb{N}$, there exists $\mathbf{\Pi}_{m+1} \notin \{\mathbf{\Pi}_m\}$, while $\mathbf{\Pi}_{m+1} \in \mathbf{\Pi}$. Thus, there exists $\{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_{m+1}\} \notin \{\boldsymbol{G}_m\}$, while $\{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_{m+1}\} \in \boldsymbol{C}$. Therefore, $\{\boldsymbol{G}_n\}$ is an open cover of $\boldsymbol{C}$ that does not have a finite subcover. Hence, $\boldsymbol{C}$ is *not compact*. □

## C.6 Proof of Proposition 3.2.2

*Proof.* Let $\boldsymbol{x} = [x^1, x^2, \ldots, x^{A+2}]^\top$ be any point in $L(\mathbf{\Pi})$ and $\epsilon > 0$, we prove that there is $\boldsymbol{y} = [y^1, y^2, \ldots, y^{A+2}]^\top \in \mathbf{\Pi} \backslash \{\boldsymbol{x}\}$ such that $d(\boldsymbol{x}, \boldsymbol{y}) < \epsilon$. Let $y^s = x^s$ for $s = 1, 2, \ldots, A$, $y^{A+1} = x^{A+1} - \delta$, and $y^{A+2} = x^{A+2} + \delta$, where $\delta < \min\left\{x^{A+1}, 1 - x^{A+2}, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{i=1}^{A+2} y^s = 1$ and $y^s > 0$ for $s = 1, 2, \ldots, A+2$ imply that $\boldsymbol{y} \in \mathbf{\Pi} \backslash \{\boldsymbol{x}\}$. Since $d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{2\delta^2} < \sqrt{\epsilon^2} = \epsilon$, $\boldsymbol{x} \in L(\mathbf{\Pi})$.

Since $\partial\mathbf{\Pi} \subset L(\mathbf{\Pi})$, every $\boldsymbol{x} \in \partial\mathbf{\Pi}$ is an element in $L(\mathbf{\Pi})$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\boldsymbol{x}, \epsilon)$ that is also an element of $\mathbf{\Pi}$. Now, consider $\boldsymbol{y} = [y^1, y^2, \ldots, y^{A+2}]^\top \in \mathcal{B}(\boldsymbol{x}, \epsilon)$ in that $y^1 = x^1 + \delta$, $y^2 = x^2 - \delta$, and $y^s = x^s$ for $s = 3, 4, \ldots, A+2$, where $\delta < \min\left\{1 - x^1, x^2, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{s=1}^{A+2} y^s = 1$, $y^s > 0$ for $s = 1, 2, \ldots, A+1$ and $y^{A+2} = 0$ imply that $\boldsymbol{y} \notin \mathbf{\Pi}$. Hence, $\boldsymbol{x} \in \partial\mathbf{\Pi}$. □

## C.7 Proof of Theorem 3.2.3

*Proof.* First of all, as proved in Appendix C.3, $d_1$ is a valid metric. Let $\mathbb{M}_{\boldsymbol{x}}$ be any point in $L(\boldsymbol{C})$, and thus, $\boldsymbol{x} = [x^1, x^2, \ldots, x^{A+2}]^\top \in L(\boldsymbol{\Pi})$. Let $\epsilon > 0$, we prove that there is $\mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C} \backslash \mathbb{M}_{\boldsymbol{x}}$ (equivalently, $\boldsymbol{y} = [y^1, y^2, \ldots, y^{A+2}]^\top \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$) such that $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon$. Let $y^s = x^s$ for $s = 1, 2, \ldots, A$, $y^{A+1} = x^{A+1} - \delta$, and $y^{A+2} = x^{A+2} + \delta$. By choosing $\delta$ satisfying the following conditions:

$$d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon \text{ and } \delta < \min\left\{x^{A+1}, 1 - x^{A+2}\right\},$$

we have $\sum_{i=1}^{A+2} y^s = 1$ and $y^s > 0$ for $s = 1, 2, \ldots, A+2$ imply that $\boldsymbol{y} \in \boldsymbol{\Pi} \backslash \{\boldsymbol{x}\}$. Therefore, we can find $\mathbb{M}_{\boldsymbol{y}} \in \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ such that $\mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C}$, and thus, $\mathbb{M}_{\boldsymbol{x}} \in L(\boldsymbol{C})$.

Since $\partial \boldsymbol{C} \subset L(\boldsymbol{C})$, every $\mathbb{M}_{\boldsymbol{x}} \in \partial \boldsymbol{C}$ is an element in $L(\boldsymbol{C})$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ that is also an element of $\boldsymbol{C}$. Now, consider $\mathbb{M}_{\boldsymbol{y}} \in \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ in that $y^1 = x^1 + \delta$ $y^2 = x^2 - \delta$, and $y^s = x^s$ for $s = 3, 4, \ldots, A+2$, where $\delta < \min\{1 - x^1, x^2\}$ and satisfies $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) < \epsilon$. Then, $y^s > 0$ for $s = 1, 2, \ldots, A+1$, and $y^{A+2} = 0$, implying that $\boldsymbol{y} \notin \boldsymbol{\Pi}$, and thus, $\mathbb{M}_{\boldsymbol{y}} \notin \boldsymbol{C}$. Hence, $\mathbb{M}_{\boldsymbol{x}} \in \partial \boldsymbol{C}$. □

## C.8 Proof of Lemma 3.2.2

*Proof.* Suppose that $(M_n^1, M_n^2, \ldots, M_n^{A+1}) \to (M^1, M^2, \ldots, M^{A+1}) \in \partial \boldsymbol{C}$ with $\sum_{s=1}^{A+1} \pi_n^s = 1$ and $1/\pi_n^s > 0$ for $s = 1, 2, \ldots, A+1$. If it were $1/\pi^1 = 0$, then $\pi_n^1 \to \infty$, but $\pi_n^1 \le \sum_{s=1}^{A+1} \pi_n^s = 1$, so that's impossible. Similarly, we cannot have $1/\pi_n^s = 0$ for any $s = 2, 3, \ldots, A+1$. Then, $\lim_{n\to\infty} \sum_{s=1}^{A+1} \pi_n^s = \sum_{s=1}^{A+1} \pi^s = 1$ so that $(M_n^1, M_n^2, \ldots, M_n^{A+1}) \in \partial \boldsymbol{C}$, and $\partial \boldsymbol{C}$ is closed. Since $\partial \boldsymbol{C}$ is a non-empty subspace of $\bar{\boldsymbol{C}}$, taking $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$, there exists a closed ball $\mathcal{B} = \mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon)$ such that $\mathcal{B} \bigcap \partial \boldsymbol{C}$ is a non-empty compact set. So the function $\mathbb{M}_{\boldsymbol{y}} \mapsto d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$ defined on $\mathcal{B} \bigcap \partial \boldsymbol{C}$ must achieve a minimum. That is, there is some $\mathbb{M}_{\boldsymbol{y}} = \mathbb{M}_{\boldsymbol{y}}^* \in \mathcal{B} \bigcap \partial \boldsymbol{C}$, which minimizes $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})$. Further, for $\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C} \backslash \mathcal{B}$, we have $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) > \epsilon \ge d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}^*)$, so it minimizes the distance on the whole of $\partial \boldsymbol{C}$. Moreover, since $\mathbb{M}_{\boldsymbol{y}}^* \notin \boldsymbol{C}$, $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}^*) > 0$. □

## C.9 Proof of Theorem 3.2.4

*Proof.* Let $\rho \in [0,1]$ and $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C}$. The affine combination of $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}$ is

$$\begin{aligned}\rho\,\mathbb{M}_{\boldsymbol{x}} + (1-\rho)\,\mathbb{M}_{\boldsymbol{y}} &= \{\boldsymbol{M} : \rho \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{rM}] + (1-\rho)\mathrm{E}_{\boldsymbol{y}}[\boldsymbol{rM}] = \rho \mathbf{1}_{A+1} + (1-\rho)\mathbf{1}_{A+1}\} \\ &= \{\boldsymbol{M} : \rho \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{rM}] + (1-\rho)\mathrm{E}_{\boldsymbol{y}}[\boldsymbol{rM}] = \mathbf{1}_{A+1}\} \in \boldsymbol{C}.\end{aligned}$$

Hence, $\boldsymbol{C}$ is *convex.*

Let $\mathbb{M}_{\boldsymbol{x}} \in \boldsymbol{C}$ with $\boldsymbol{x} \in \boldsymbol{\Pi}$. There is $\eta > 0$ such that

$$\eta = \min_{\mathbb{M}_{\boldsymbol{y}} \in \partial \boldsymbol{C}} \{d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}})\},$$

where $\boldsymbol{y} \in \partial\boldsymbol{\Pi}$. Then, since $\boldsymbol{C} = \bar{\boldsymbol{C}} \backslash \partial \boldsymbol{C}$, by choosing $\epsilon < \eta$, we have $\mathcal{B}(\mathbb{M}_{\boldsymbol{x}}, \epsilon) \subset \boldsymbol{C}$. Hence, $\boldsymbol{C}$ is *open* in $\left(\bar{\boldsymbol{C}}, d_1\right)$.

Next, we prove that $\boldsymbol{C}$ is bounded in $\left(\bar{\boldsymbol{C}}, d_1\right)$. First, let $\boldsymbol{\kappa_v}$ and $\boldsymbol{\kappa_u}$ denote the coefficient vectors of $\boldsymbol{v}(\boldsymbol{x})$ and $\boldsymbol{u}(\boldsymbol{x})$, respectively, for $\boldsymbol{x} \in \bar{\boldsymbol{\Pi}}$ such that

$$\begin{cases} \boldsymbol{\kappa_v} = \begin{bmatrix} (\boldsymbol{r}')^{-1}\,\mathbf{1}_{A+1} \\ 0 \end{bmatrix} \text{ and } \boldsymbol{\kappa_u} = \begin{bmatrix} -(\boldsymbol{r}')^{-1}(\boldsymbol{r}'') \\ 1 \end{bmatrix}, & \boldsymbol{x} \in \boldsymbol{\Pi} \\ \boldsymbol{\kappa_v} = (\boldsymbol{r}')^{-1}\,\mathbf{1}_{A+1} \text{ and } \boldsymbol{\kappa_u} = \mathbf{0}_{A+1}, & \boldsymbol{x} \in \partial\boldsymbol{\Pi}. \end{cases}$$

Given that $\boldsymbol{w} \in W(\boldsymbol{x}, \boldsymbol{y})$, we have

$$\begin{aligned}\mathcal{D} &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} d_2(M^{s_{\boldsymbol{x}}}, M^{s_{\boldsymbol{y}}}) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \left(|\kappa_{\boldsymbol{v}}^{s_{\boldsymbol{x}}}/x^{s_{\boldsymbol{x}}} - \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}}/y^{s_{\boldsymbol{y}}}| + |\kappa_{\boldsymbol{u}}^{s_{\boldsymbol{x}}}/x^{s_{\boldsymbol{x}}} - \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{y}}}/y^{s_{\boldsymbol{y}}}|\right) \\ &= \sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} \sum_{s_{\boldsymbol{y}}=1}^{S_{\boldsymbol{y}}} \left(|w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_i^{\boldsymbol{v}} - w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}}| + |w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{x}}} - w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{u}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}}|\right)\end{aligned}$$

Since $\sum_{s_{\boldsymbol{x}}=1}^{S_{\boldsymbol{x}}} w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} = y^{s_{\boldsymbol{y}}}$, for every $s_{\boldsymbol{x}} = 1, 2, \ldots, S_{\boldsymbol{x}}, s_{\boldsymbol{y}} = 1, 2, \ldots, S_{\boldsymbol{y}}$, $w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} x^{s_{\boldsymbol{x}}} \in [0, y^{s_{\boldsymbol{y}}}]$, and thus, $w_{s_{\boldsymbol{y}} s_{\boldsymbol{x}}} \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}} x^{s_{\boldsymbol{x}}}/y^{s_{\boldsymbol{y}}} \in [0, \kappa_{\boldsymbol{v}}^{s_{\boldsymbol{y}}}]$. Therefore, $\mathcal{D}$ is bounded, implying that there is $0 < R < \infty$ such that $d_1(\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}}) \le R$ for all $\mathbb{M}_{\boldsymbol{x}}, \mathbb{M}_{\boldsymbol{y}} \in \boldsymbol{C}$. Hence, $\boldsymbol{C}$ is *bounded.*

Last, to show that $\boldsymbol{C}$ is not compact, we just need one example of an open cover that has no finite open subcovers. Let $\{\boldsymbol{G}_n\} = \{\mathbb{M}_{\boldsymbol{\pi}} \,|\, \boldsymbol{\pi} \in \boldsymbol{\Pi}_n, n \in \mathbb{N}\}$, where

$$\{\boldsymbol{\Pi}_n\} = \left\{ \left[\pi^1, \pi^2, \ldots, \pi^{A+1}, 1 - \sum_{s=1}^{A+1} \pi^s\right]^{\top} \in (\mathbb{R})_{++}^{A+2} : \sum_{s=1}^{A+1} \pi^s \in \left(\frac{1}{n}, \frac{n-1}{n}\right), n \in \mathbb{N} \right\}.$$

Notice that, if this gives us an invalid segment such as $(1,0)$, we treat it as an empty element. Here, for any $b = \sum_{s=1}^{A+1} \pi^s \in (0,1)$, the Archimedean Property provides an $n \in \mathbb{N}$ such that $n > \max\left\{\frac{1}{b}, \frac{1}{1-b}\right\}$. Then,

$$\begin{aligned} &nb > 1 \text{ and } n - nb > 1 \\ \Rightarrow\ &1 < nb < n-1 \\ \Rightarrow\ &b \in \left(\frac{1}{n}, \frac{n}{n-1}\right). \end{aligned}$$

Thus, every element of $\mathbf{\Pi}$ is in $\{\mathbf{\Pi}_n\}$ for some $n \in \mathbb{N}$, and therefore, every element of $\boldsymbol{C}$ is in $\{\boldsymbol{G}_n\}$ for some $n \in \mathbb{N}$, suggesting that $\boldsymbol{C} \subset \bigcup_{n=1}^{\infty}\{\boldsymbol{G}_n\}$. Moreover, since for any $n \in \mathbb{N}$, $\boldsymbol{G}_n$ has a neighbourhood centred at $\boldsymbol{G}_n$ with radius $\epsilon > 0$ lying in the set, $\{\boldsymbol{G}_n\}$ is an open cover of $\boldsymbol{C}$. Let $k, l \in \mathbb{N}$ such that $k > l > 2$, we have

$$\frac{1}{k} < \frac{1}{l} < 1 - \frac{1}{l} < 1 - \frac{1}{k} \Rightarrow \{\mathbf{\Pi}_l\} \subset \{\mathbf{\Pi}_k\} \Rightarrow \{\boldsymbol{G}_l\} \subset \{\boldsymbol{G}_k\}.$$

Therefore, for any finite $m \in \mathbb{N}$, $\bigcup_{n=1}^{m}\{\boldsymbol{G}_n\} = \{\boldsymbol{G}_m\} = \{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_m, m \in \mathbb{N}\}$. However, for any $m \in \mathbb{N}$, there exists $\mathbf{\Pi}_{m+1} \notin \{\mathbf{\Pi}_m\}$, while $\mathbf{\Pi}_{m+1} \in \mathbf{\Pi}$. Thus, there exists $\{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_{m+1}\} \notin \{\boldsymbol{G}_m\}$, while $\{\mathbb{M}_{\boldsymbol{\pi}} | \boldsymbol{\pi} \in \mathbf{\Pi}_{m+1}\} \in \boldsymbol{C}$. Therefore, $\{\boldsymbol{G}_n\}$ is an open cover of $\boldsymbol{C}$ that does not have a finite subcover. Hence, $\boldsymbol{C}$ is *not compact*. □

## C.10 Proof of Proposition 3.2.4

*Proof.* We first prove by induction that, for all $n \in \mathbb{Z}_+$ and $S = A + 1 + n$, there is $\boldsymbol{x}_t = [x_t^1, x_t^2, \ldots, x_t^S]^\top \in L(\mathbf{\Pi}_t)$ and $\boldsymbol{y}_t = [y_t^1, y_t^2, \ldots, y_t^S]^\top \in \mathbf{\Pi}_t \backslash \{\boldsymbol{x}_t\}$ such that $d(\boldsymbol{x}_t, \boldsymbol{y}_t) < \epsilon$ for $\epsilon > 0$.

**Base case:** When $n = 1$, $S = A + 2$, let $\boldsymbol{x}_t = [x_t^1, x_t^2, \ldots, x_t^{A+2}]^\top$ be any point in $L(\mathbf{\Pi}_t)$. Let $\epsilon > 0$ and $\boldsymbol{y}_t = [y_t^1, y_t^2, \ldots, y_t^S]^\top \in \mathbf{\Pi}_t \backslash \{\boldsymbol{x}_t\}$ such that $y_t^s = x_t^s$ for $s = 1, 2, \ldots, A$, $y_t^{A+1} = x_t^{A+1} - \delta$, and $y_t^{A+2} = x_t^{A+2} + \delta$ and $\delta < \min\left\{x_t^{A+1}, 1 - x_t^{A+2}, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{i=1}^{A+2} y_t^s = 1$ and $y_t^s > 0$ for $a = 1, 2, \ldots, A+2$, implying that $\boldsymbol{y}_t \in \mathbf{\Pi}_t \backslash \{\boldsymbol{x}_t\}$. Since $d(\boldsymbol{x}_t, \boldsymbol{y}_t) = \sqrt{2\delta^2} < \sqrt{\epsilon^2} = \epsilon$, $\boldsymbol{x}_t \in L(\mathbf{\Pi}_t)$.

**Induction step:** Let $k \in \mathbb{Z}_+$ be given and suppose our statement is true for $n = k$. Then, if $x_t^{A+k+2} = \eta > 0$, holding all other elements in $\boldsymbol{x}_t(n = k)$ and $\boldsymbol{y}_t(n = k)$ fixed, for any $i < A+k+2$ with $x_t^i(n = k) > \eta$[1], given $x_t^i(n = k + 1) = x_t^i(n = k) - \eta$, there is $y_t^i(n = k + 1) = y_t^i(n = k) - \eta$

[1]We use $x_t^i(n = k)$ to denote the $i^{th}$ element in $\boldsymbol{x}_t$ for $n = k$.

such that

$$\left(x_t^i(n=k+1) - y_t^i(n=k+1)\right)^2 = \left(x_t^i(n=k) - \gamma - y_t^i(n=k) + \gamma\right)^2$$
$$= \left(x_t^i(n=k) - y_t^i(n=k)\right)^2$$

and $\left(x_t^{A+k+2}(n=k+1) - y_t^{A+k+2}(n=k+1)\right)^2 = 0$. Therefore, $d(\boldsymbol{x}_t(n=k+1), \boldsymbol{y}_t(n=k+1)) = d(\boldsymbol{x}_t(n=k), \boldsymbol{y}_t(n=k)) < \epsilon$. If $x_t^{A+k+2} = 0$, for any $x_t^i(n=k) > y_t^i(n=k)$, holding all other elements in $\boldsymbol{y}_t(n=k)$ fixed, let $y_t^{A+k+2}(n=k+1) = \gamma < x_t^i(n=k) - y_t^i(n=k)$ and $y_t^i(n=k+1) = y_t^i(n=k) - \gamma$, then

$$\begin{aligned}
&\left(x_t^i(n=k+1) - y_t^i(n=k+1)\right)^2 + \left(x_t^{A+k+2}(n=k+1) - y_t^{A+k+2}(n=k+1)\right)^2 \\
&= \left(x_t^i(n=k) - y_t^i(n=k) + \gamma\right)^2 + \gamma^2 \\
&= \left(x_t^i(n=k) - y_t^i(n=k)\right)^2 - 2\gamma\left(x_t^i(n=k) - y_t^i(n=k)\right) + 2\gamma^2 \\
&< \left(x_t^i(n=k) - y_t^i(n=k)\right)^2.
\end{aligned}$$

Therefore, $d(\boldsymbol{x}_t(n=k+1), \boldsymbol{y}_t(n=k+1)) < d(\boldsymbol{x}_t(n=k), \boldsymbol{y}_t(n=k)) < \epsilon$.
**Conclusion:** By the principal of induction, for all $n \in \mathbb{Z}_+$ and $S = A+1+n$, there is $\boldsymbol{x}_t = [x_t^1, x_t^2, \dots, x_t^S]^\top \in L(\boldsymbol{\Pi}_t)$ and $\boldsymbol{y}_t = [y_t^1, y_t^2, \dots, y_t^S]^\top \in \boldsymbol{\Pi}_t \backslash \{\boldsymbol{x}_t\}$ such that $d(\boldsymbol{x}_t, \boldsymbol{y}_t) < \epsilon$ for $\epsilon > 0$.

Since $\partial\boldsymbol{\Pi}_t \subset L(\boldsymbol{\Pi}_t)$, every $\boldsymbol{x}_t \in \partial\boldsymbol{\Pi}_t$ is an element in $L(\boldsymbol{\Pi}_t)$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\boldsymbol{x}_t, \epsilon)$ that is also an element of $\boldsymbol{\Pi}_t$. Now, consider $\boldsymbol{y}_t = [y_t^1, y_t^2, \dots, y_t^S]^\top \in \mathcal{B}(\boldsymbol{x}_t, \epsilon)$ in that $y_t^1 = x_t^1 + \delta$, $y_t^2 = x_t^2 - \delta$, and $y_t^s = x_t^s$ for $s = 3, 4, \dots, S$, where $\delta < \min\left\{1 - x_t^1, x_t^2, \epsilon/\sqrt{2}\right\}$. Then, $\sum_{s=1}^S y_t^s = 1$, $y_t^s > 0$ for $s = 1, 2, \dots, A+1$, and $y_t^s = 0$ for $s = A+2, A+3, \dots, S$ imply that $\boldsymbol{y}_t \notin \boldsymbol{\Pi}_t$. Hence, $\boldsymbol{x}_t \in \partial\boldsymbol{\Pi}_t$. □

## C.11 Proof of Theorem 3.2.7

First notice that, similar to the proof in the 1-1-2-3 case, $d_1$ is a valid metric. Then, we prove by induction that, for all $n \in \mathbb{Z}_+$ and $S = A+1+n$, let $\mathbb{m}_{\boldsymbol{x}}$ be any point in $L(\boldsymbol{c}_t)$ and $\epsilon > 0$, we want to prove that there is $\mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t \backslash \mathbb{m}_{\boldsymbol{x}_t}$ such that $d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}) < \epsilon$, where $\boldsymbol{y}_t \in \boldsymbol{\Pi}_t \backslash \{\boldsymbol{x}_t\}$.
**Base case:** Since $\mathbb{m}_{\boldsymbol{x}} \in L(\boldsymbol{c}_t)$, $\boldsymbol{x}_t \in L(\boldsymbol{\Pi}_t)$. Let $y_t^s = x_t^s$ for $s = 1, 2, \dots, A$, $y_t^{A+1} = x_t^{A+1} - \delta$, and $y_t^{A+2} = x_t^{A+2} + \delta$. By choosing $\delta$ satisfying the following conditions:

$$d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}) < \epsilon \text{ and } \delta < \min\left\{x_t^{A+1}, 1 - x_t^{A+2}\right\},$$

so that we have $\sum_{i=1}^{A+2} y_t^s = 1$ and $y_t^s > 0$ for $s = 1, 2, \ldots, A+2$ imply that $\boldsymbol{y}_t \in \boldsymbol{\Pi}_t \backslash \{\boldsymbol{x}_t\}$. Therefore, we can find $\mathbb{m}_{\boldsymbol{y}_t} \in \mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon)$ such that $\mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t$, and thus, $\mathbb{m}_{\boldsymbol{x}_t} \in L(\boldsymbol{c}_t)$.
**Induction step:** Let $k \in \mathbb{Z}_+$ be given and suppose our statement is true for $n = k$. Then, for every $\boldsymbol{x}_t(n = k+1)$ such that $x_t^{A+k+2} = \eta > 0$ and $x_t^i(n = k+1) = x_t^i(n = k) - \lambda_i \eta$ for $i = 1, 2, \ldots, A+k+1$, where $\sum_{i=1}^{A+1} \lambda_i = 1$ and $\lambda_i > 0$. By setting $y_t^{A+k+2} = \eta$ and $y_t^i(n = k+1) = y_t^i(n = k) - \lambda_i \eta > 0$ $i = 1, 2, \ldots, A+k+1$, we have $d_1(\mathbb{m}_{\boldsymbol{x}_t(n=k+1)}, \mathbb{m}_{\boldsymbol{y}_t(n=k+1)}) = d_1(\mathbb{m}_{\boldsymbol{x}_t(n=k)}, \mathbb{m}_{\boldsymbol{y}_t(n=k)}) < \epsilon$. If $x_t^{A+k+2} = 0$, by choosing $\eta$ and $\lambda_i \geq 0$ $i = 1, 2, \ldots, A+k+1$ such that $y_t^{A+k+2} = \eta > 0$, $y_t^i(n = k+1) = y_t^i(n = k) - \lambda_i \eta > 0$, and $\sum_{i=1}^{A+1} \lambda_i = 1$, and satisfies the condition that $d_1(\mathbb{m}_{\boldsymbol{x}_t(n=k+1)}, \mathbb{m}_{\boldsymbol{y}_t(n=k+1)}) < \epsilon$ so that we have $\sum_{i=1}^{A+k+2} y_t^s = 1$ and $y_t^s > 0$ for $s = 1, 2, \ldots, A+k+2$ imply that $\boldsymbol{y}_t \in \boldsymbol{\Pi}_t \backslash \{\boldsymbol{x}_t\}$. Therefore, we can find $\mathbb{m}_{\boldsymbol{y}_t} \in \mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon)$ such that $\mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t$, and thus, $\mathbb{m}_{\boldsymbol{x}_t} \in L(\boldsymbol{c}_t)$.

Since $\partial \boldsymbol{c}_t \subset L(\boldsymbol{c}_t)$, every $\mathbb{m}_{\boldsymbol{x}_t} \in \partial \boldsymbol{c}_t$ is an element in $L(\boldsymbol{c}_t)$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon)$ that is also an element of $\boldsymbol{c}_t$. Now, consider $\mathbb{m}_{\boldsymbol{y}_t} \in \mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon)$ in that $y_t^1 = x_t^1 + \delta$ $y_t^2 = x_t^2 - \delta$, and $y_t^s = x_t^s$ for $s = 3, 4, \ldots, A+2$, where $\delta < \min\{1 - x_t^1, x_t^2\}$ and satisfies $d_1(\mathbb{m}_{\boldsymbol{x}}, \mathbb{m}_{\boldsymbol{y}}) < \epsilon$. Then, $y_t^s > 0$ for $s = 1, 2, \ldots, A+1$, and $y_t^s = 0$ for $s = A+2, A+3, \ldots, S$, implying that $\boldsymbol{y}_t \notin \boldsymbol{\Pi}_t$, and thus, $\mathbb{m}_{\boldsymbol{y}_t} \notin \boldsymbol{c}_t$. Hence, $\mathbb{m}_{\boldsymbol{x}_t} \in \partial \boldsymbol{c}_t$.

## C.12 Proof of Lemma 3.2.3

*Proof.* Suppose that $(m_{t,n}^1, m_{t,n}^2, \ldots, m_{t,n}^{A+1}) \to (m_t^1, m_t^2, \ldots, m_t^{A+1}) \in \partial \boldsymbol{c}_t$ with $\sum_{s=1}^{A+1} \pi_{t,n}^s = 1$ and $1/\pi_{t,n}^s > 0$ for $s = 1, 2, \ldots, A+1$. If it were $1/\pi_t^1 = 0$, then $\pi_{t,n}^1 \to \infty$, but $\pi_{t,n}^1 \leq \sum_{s=1}^{A+1} \pi_{t,n}^s = 1$, so that's impossible. Similarly, we cannot have $1/\pi_{t,n}^s = 0$ for any $s = 2, 3, \ldots, A+1$. Then, $\lim_{n\to\infty} \sum_{s=1}^{A+1} \pi_{t,n}^s = \sum_{s=1}^{A+1} \pi_t^s = 1$ so that $(m_{t,n}^1, m_{t,n}^2, \ldots, m_{t,n}^{A+1}) \in \partial \boldsymbol{c}_t$, and $\partial \boldsymbol{c}_t$ is closed. Since $\partial \boldsymbol{c}_t$ is a non-empty subspace of $\bar{\boldsymbol{c}}_t$, taking $\mathbb{m}_{\boldsymbol{x}_t} \in \boldsymbol{c}_t$, there exists a closed ball $\mathcal{B} = \mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon)$ such that $\mathcal{B} \bigcap \partial \boldsymbol{c}_t$ is a non-empty compact set. So the function $\mathbb{m}_{\boldsymbol{y}_t} \mapsto d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t})$ defined on $\mathcal{B} \bigcap \partial \boldsymbol{c}_t$ must achieve a minimum. That is, there is some $\mathbb{m}_{\boldsymbol{y}_t} = \mathbb{m}_{\boldsymbol{y}}^* \in \mathcal{B} \bigcap \partial \boldsymbol{c}_t$, which minimizes $d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t})$. Further, for $\mathbb{m}_{\boldsymbol{y}_t} \in \partial \boldsymbol{c}_t \backslash \mathcal{B}$, we have $d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}) > \epsilon \geq d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}^*)$, so it minimizes the distance on the whole of $\partial \boldsymbol{c}_t$. Moreover, since $\mathbb{m}_{\boldsymbol{y}_t}^* \notin \boldsymbol{c}_t$, $d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}^*) > 0$. □

## C.13 Proof of Theorem 3.2.8

*Proof.* Let $\rho \in [0,1]$ and $\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t$. The affine combination of $\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}$ is

$$\begin{aligned}\rho\,\mathbb{m}_{\boldsymbol{x}_t} + (1-\rho)\,\mathbb{m}_{\boldsymbol{y}_t} &= \left\{\boldsymbol{m}_t : \rho\mathrm{E}_{\boldsymbol{x}_t}[\boldsymbol{r}_t\boldsymbol{m}_t] + (1-\rho)\mathrm{E}_{\boldsymbol{y}_t}[\boldsymbol{r}_1\boldsymbol{m}_t] = \rho\mathbf{1}_{A+1} + (1-\rho)\mathbf{1}_{A+1}\right\}\\ &= \left\{\boldsymbol{m}_t : \rho\mathrm{E}_{\boldsymbol{x}_t}[\boldsymbol{r}_t\boldsymbol{m}_t] + (1-\rho)\mathrm{E}_{\boldsymbol{y}_t}[\boldsymbol{r}_t\boldsymbol{m}_t] = \mathbf{1}_{A+1}\right\} \in \boldsymbol{c}_t.\end{aligned}$$

Hence, $\boldsymbol{c}_t$ is *convex.*

Then let $\mathbb{m}_{\boldsymbol{x}_t} \in \boldsymbol{c}_t$ with $\boldsymbol{x}_t \in \boldsymbol{\Pi}_t$. There is $\eta > 0$ such that

$$\eta = \min_{\mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t} \left\{d_1(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t})\right\},$$

where $\boldsymbol{y}_t \in \partial\boldsymbol{\Pi}_t$. Then, since $\boldsymbol{c}_t = \bar{\boldsymbol{c}}_t \backslash \partial\boldsymbol{c}_t$, by choosing $\epsilon < \eta$, we have $\mathcal{B}(\mathbb{m}_{\boldsymbol{x}_t}, \epsilon) \subset \boldsymbol{c}_t$. Hence, $\boldsymbol{c}_t$ is *open* in $(\bar{\boldsymbol{c}}_t, d_1)$.

Next, we prove that $\boldsymbol{c}_t$ is bounded in $(\bar{\boldsymbol{c}}_t, d_1)$. First, let $\boldsymbol{\kappa}_{\boldsymbol{v}_t}$ and $\boldsymbol{\kappa}_{\boldsymbol{u}_t}$ denote the coefficient vectors of $\boldsymbol{v}_t(\boldsymbol{x}_t)$ and $\boldsymbol{u}_t(\boldsymbol{x}_t)$, respectively, for $\boldsymbol{x}_t \in \bar{\boldsymbol{\Pi}}_t$. Given that $\boldsymbol{w}_t \in W(\boldsymbol{x}_t, \boldsymbol{y}_t)$, we have

$$\begin{aligned}\mathcal{D} &= \sum_{s_{\boldsymbol{x}_t}=1}^{S_{\boldsymbol{x}_t}} \sum_{s_{\boldsymbol{y}_t}=1}^{S_{\boldsymbol{y}_t}} w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}} x_t^{s_{\boldsymbol{x}_t}} d_2(m_t^{s_{\boldsymbol{x}_t}}, m_t^{s_{\boldsymbol{y}_t}})\\ &= \sum_{s_{\boldsymbol{x}_t}=1}^{S_{\boldsymbol{x}_t}} \sum_{s_{\boldsymbol{y}_t}=1}^{S_{\boldsymbol{y}_t}} w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}} x_t^{s_{\boldsymbol{x}_t}} \left(|\kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{x}_t}}/x_t^{s_{\boldsymbol{x}_t}} - \kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{y}_t}}/y_t^{s_{\boldsymbol{y}_t}}| + |\kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{x}_t}}/x_t^{s_{\boldsymbol{x}_t}} - \kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{y}_t}}/y_t^{s_{\boldsymbol{y}_t}}|\right)\\ &= \sum_{s_{\boldsymbol{x}_t}=1}^{S_{\boldsymbol{x}_t}} \sum_{s_{\boldsymbol{y}_t}=1}^{S_{\boldsymbol{y}_t}} \left(|w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{x}_t}} - w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{y}_t}} x_t^{s_{\boldsymbol{x}_t}}/y_t^{s_{\boldsymbol{y}_t}}| + |w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{x}_t}} - w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{y}_t}} x_t^{s_{\boldsymbol{x}_t}}/y_t^{s_{\boldsymbol{y}_t}}|\right).\end{aligned}$$

Since $\sum_{s_{\boldsymbol{x}_t}=1}^{S_{\boldsymbol{x}_t}} w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}} x_t^{s_{\boldsymbol{x}_t}} = y_t^{s_{\boldsymbol{y}_t}}$, for every $s_{\boldsymbol{x}_t} = 1, 2, \ldots, S_{\boldsymbol{x}_t}$, $s_{\boldsymbol{y}_t} = 1, 2, \ldots, S_{\boldsymbol{y}_t}$, $w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}} x_t^{s_{\boldsymbol{x}_t}} \in [0, y_t^{s_{\boldsymbol{y}_t}}]$. Thus, $w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{y}_t}} x_t^{s_{\boldsymbol{x}_t}}/y_t^{s_{\boldsymbol{y}_t}} \in [0, \kappa_{\boldsymbol{v}_t}^{s_{\boldsymbol{y}_t}}]$ and $w_{s_{\boldsymbol{y}_t}s_{\boldsymbol{x}_t}}\kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{y}_t}} x_t^{s_{\boldsymbol{x}_t}}/y_t^{s_{\boldsymbol{y}_t}} \in [0, \kappa_{\boldsymbol{u}_t}^{s_{\boldsymbol{y}_t}}]$. Therefore, $\mathcal{D}$ is bounded, implying that there is $0 < R < \infty$ such that $d_1\left(\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t}\right) \le R$ for all $\mathbb{m}_{\boldsymbol{x}_t}, \mathbb{m}_{\boldsymbol{y}_t} \in \boldsymbol{c}_t$. Hence, $\boldsymbol{c}_t$ is *bounded.*

Last, to show that $\boldsymbol{c}_t$ is not compact, we just need one example of an open cover that has no finite open subcovers. Let $\{\boldsymbol{G}_{t,n}\} = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \boldsymbol{\Pi}_{t,n}, n \in \mathbb{N}\}$, where

$$\{\boldsymbol{\Pi}_{t,n}\} = \left\{\left[\pi_t^1, \pi_t^2, ..., \pi_t^{S-1}, 1 - \sum_{s=1}^{S-1}\pi_t^s\right]^\top \in (\mathbb{R})_{++}^S : \sum_{s=1}^{S-1}\pi_t^s \in \left(\frac{1}{n}, \frac{n-1}{n}\right), n \in \mathbb{N}\right\}.$$

Notice that, if this gives us an invalid segment such as $(1,0)$, we treat it as an empty element. Here, for any $b = \sum_{s=1}^{S-1} \pi_t^s \in (0,1)$, the Archimedean Property provides an $n \in \mathbb{N}$ such that $n > \max\left\{\frac{1}{b}, \frac{1}{1-b}\right\}$. Then,

$$\begin{aligned} &nb > 1 \text{ and } n - nb > 1 \\ \Rightarrow\, &1 < nb < n - 1 \\ \Rightarrow\, &b \in \left(\frac{1}{n}, \frac{n}{n-1}\right). \end{aligned}$$

Thus, every element of $\mathbf{\Pi}_t$ is in $\{\mathbf{\Pi}_{t,n}\}$ for some $n \in \mathbb{N}$, and therefore, every element of $\boldsymbol{c}_t$ is in $\{\boldsymbol{G}_{t,n}\}$ for some $n \in \mathbb{N}$, suggesting that $\boldsymbol{c}_t \subset \bigcup_{n=1}^{\infty} \{\boldsymbol{G}_{t,n}\}$. Moreover, since for any $n \in \mathbb{N}$, $\boldsymbol{G}_{t,n}$ has a neighbourhood centred at $\boldsymbol{G}_{t,n}$ with radius $\epsilon > 0$ lying in the set, $\{\boldsymbol{G}_{t,n}\}$ is an open cover of $\boldsymbol{c}_t$. Let $k, l \in \mathbb{N}$ such that $k > l > 2$, we have

$$\frac{1}{k} < \frac{1}{l} < 1 - \frac{1}{l} < 1 - \frac{1}{k} \Rightarrow \{\mathbf{\Pi}_{t,l}\} \subset \{\mathbf{\Pi}_{t,k}\} \Rightarrow \{\boldsymbol{G}_{t,l}\} \subset \{\boldsymbol{G}_{t,k}\}.$$

Therefore, for any finite $m \in \mathbb{N}$, $\bigcup_{n=1}^{m} \{\boldsymbol{G}_{t,n}\} = \{\boldsymbol{G}_{t,m}\} = \{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \mathbf{\Pi}_{t,m}, m \in \mathbb{N}\}$. However, for any $m \in \mathbb{N}$, there exists $\mathbf{\Pi}_{t,m+1} \notin \{\mathbf{\Pi}_{t,m}\}$, while $\mathbf{\Pi}_{t,m+1} \in \mathbf{\Pi}_t$. Thus, there exists $\{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \mathbf{\Pi}_{t,m+1}\} \notin \{\boldsymbol{G}_{t,m}\}$, while $\{\mathbb{m}_{\boldsymbol{\pi}_t}, \boldsymbol{\pi}_t \in \mathbf{\Pi}_{t,m+1}\} \in \boldsymbol{c}_t$. Therefore, $\{\boldsymbol{G}_{t,n}\}$ is an open cover of $\boldsymbol{c}_t$ that does not have a finite subcover. Hence, $\boldsymbol{c}_t$ is *not compact*. □

## C.14 Proof of Proposition 3.3.1

*Proof.* Let $\boldsymbol{\phi}_t = (\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx))$ be any point in $L(\boldsymbol{\Phi}_t)$. Let $\epsilon > 0$, we want to prove that there is $\boldsymbol{\phi}_t' = \left(\boldsymbol{\mu}_t^{B'}, \boldsymbol{\sigma}_t^{B'}, \boldsymbol{\mu}_t^{J'}, \boldsymbol{\sigma}_t^{J'}, \boldsymbol{v}'(dx)\right) \in \boldsymbol{\Phi}_t \backslash \{\boldsymbol{\phi}_t\}$ such that $d(\boldsymbol{\phi}_t, \boldsymbol{\phi}_t') < \epsilon$. Let $\boldsymbol{\mu}_t^{B'} = \boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^{B'} = \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^{J'} = \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^{J'} = \boldsymbol{\sigma}_t^J$, and $\boldsymbol{v}'(dx) = \boldsymbol{v}_t(dx) + \delta$, whence Equation (3.19) and 3.20 hold. Then, we have $\boldsymbol{v}'(dx) > \mathbf{0}_A$, implying that $\boldsymbol{\phi}_t' \in \boldsymbol{\Phi}_t \backslash \{\boldsymbol{\phi}_t\}$ and $d(\boldsymbol{\phi}_t, \boldsymbol{\phi}_t') < \epsilon$. Hence, $\boldsymbol{\phi}_t \in L(\boldsymbol{\Phi}_t)$.

Since $\partial \boldsymbol{\Phi}_t \subset L(\boldsymbol{\Phi}_t)$, every $\boldsymbol{\phi}_t \in \partial \boldsymbol{\phi}_t$ is an element in $L(\boldsymbol{\Phi}_t)$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\boldsymbol{\Phi}_t, \epsilon)$ that is also an element of $\boldsymbol{\phi}_t$. Now, consider $\boldsymbol{\phi}_t' = (\boldsymbol{\mu}_t^{B'}, \boldsymbol{\sigma}_t^{B'}, \boldsymbol{\mu}_t^{J'}, \boldsymbol{\sigma}_t^{J'}, \boldsymbol{v}'(dx)) \in \mathcal{B}(\boldsymbol{\Phi}_t, \epsilon)$ in that $d(\boldsymbol{\mu}_t^B, \boldsymbol{\mu}_t^{B'}) < \epsilon$, $\boldsymbol{\sigma}_t^{B'} = \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^{J'} = \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^{J'} = \boldsymbol{\sigma}_t^J$, $\boldsymbol{v}'(dx) = \boldsymbol{v}_t(dx)$, and Equation (3.20) holds. Then, $\boldsymbol{v}_t(dx) = 0$, implying that $\boldsymbol{\phi}_t \notin \boldsymbol{\Phi}_t$. Hence, $\boldsymbol{\phi}_t \in \partial \boldsymbol{\Phi}_t$. □

## C.15 Proof of Theorem 3.3.1

*Proof.* Let $\boldsymbol{c}(\boldsymbol{\phi}_t)$ be any point in $L(\boldsymbol{C}_t)$, and thus, $\boldsymbol{\phi}_t = (\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)) \in L(\boldsymbol{\Phi}_t)$. Let $\epsilon > 0$, we want to prove that there is $\mathbb{m}(\boldsymbol{\phi}_t') \in \boldsymbol{c}_t \backslash \mathbb{m}(\boldsymbol{\phi}_t)$ such that $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')) < \epsilon$, where $\boldsymbol{\phi}_t' = \left(\boldsymbol{\mu}_t^{B'}, \boldsymbol{\sigma}_t^{B'}, \boldsymbol{\mu}_t^{J'}, \boldsymbol{\sigma}_t^{J'}, \boldsymbol{v}_t'(dx)\right) \in \boldsymbol{\Phi}_t \backslash \{\boldsymbol{\phi}_t\}$. Let $\boldsymbol{\mu}_t^{B'} = \boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^{B'} = \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^{J'} = \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^{J'} = \boldsymbol{\sigma}_t^J$, and $\boldsymbol{v}_t'(dx) = \boldsymbol{v}_t(dx) + \boldsymbol{\delta}$. By choosing $\boldsymbol{\delta}$ satisfying $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')) < \epsilon$, and Equation (3.19) and 3.20 hold, we have $\boldsymbol{v}_t'(dx) > \boldsymbol{0}_A$ implying that $\boldsymbol{\phi}_t' \in \boldsymbol{\Phi}_t \backslash \{\boldsymbol{\phi}_t\}$. Therefore, $\mathbb{m}(\boldsymbol{\phi}_t') \in \mathcal{B}(\mathbb{m}(\boldsymbol{\phi}_t), \epsilon)$ such that $\mathbb{m}(\boldsymbol{\phi}_t') \in \boldsymbol{c}_t \backslash \mathbb{m}(\boldsymbol{\phi}_t)$, and thus, $\mathbb{m}(\boldsymbol{\phi}_t) \in L(\boldsymbol{c}_t)$.

Since $\partial \boldsymbol{c}_t \subset L(\boldsymbol{c}_t)$, every $\mathbb{m}(\boldsymbol{\phi}_t) \in \partial \boldsymbol{c}_t$ is an element in $L(\boldsymbol{c}_t)$. Therefore, for $\epsilon > 0$, there is at least one point in $\mathcal{B}(\mathbb{m}(\boldsymbol{\phi}_t), \epsilon)$ that is also an element of $\boldsymbol{c}_t$. Now, consider $\mathbb{m}(\boldsymbol{\phi}_t') \in \mathcal{B}(\mathbb{m}(\boldsymbol{\phi}_t), \epsilon)$ in that $|\boldsymbol{\mu}_t^B - \boldsymbol{\mu}_t^{B'}| < \boldsymbol{\delta}$, $\boldsymbol{\sigma}_t^{B'} = \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^{J'} = \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^{J'} = \boldsymbol{\sigma}_t^J$, $\boldsymbol{v}_t'(dx) = \boldsymbol{v}_t(dx)$, and $\boldsymbol{\delta}$ is chosen to satisfy $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}_t(\boldsymbol{\phi}_t')) < \epsilon$ and Equation (3.20) holds. Then, $\boldsymbol{v}'(dx) = 0$, implying that $\boldsymbol{\phi}_t' \notin \boldsymbol{\Phi}_t$, and thus, $\mathbb{m}_t(\boldsymbol{\phi}_t') \notin \boldsymbol{c}_t$. Hence, $\mathbb{m}(\boldsymbol{\phi}_t) \in \partial \boldsymbol{c}_t$. □

## C.16 Proof of Lemma 3.3.1

Since $\partial \boldsymbol{c}_t$ is a non-empty closed subspace of $\bar{\boldsymbol{c}}_t$, taking $\mathbb{m}(\boldsymbol{\phi}_t) \in \boldsymbol{c}_t$, there exists a closed ball $\mathcal{B} = \mathcal{B}(\mathbb{m}(\boldsymbol{\phi}_t), \epsilon)$ such that $\mathcal{B} \bigcap \partial \boldsymbol{c}_t$ is a non-empty compact set. So the function $\mathbb{m}(\boldsymbol{\phi}_t') \mapsto d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t'))$ defined on $\mathcal{B} \bigcap \partial \boldsymbol{c}_t$ must achieve a minimum. That is, there is some $\mathbb{m}(\boldsymbol{\phi}_t') = \mathbb{m}(\boldsymbol{\phi}_t^*) \in \mathcal{B} \bigcap \partial \boldsymbol{c}_t$, which minimizes $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t'))$. Further, for $\mathbb{m}(\boldsymbol{\phi}_t') \in \partial \boldsymbol{c}_t \backslash \mathcal{B}$, we have $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')) > \epsilon \geq d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t^*))$, so it minimizes the distance on the whole of $\partial \boldsymbol{c}_t$. Moreover, since $\mathbb{m}(\boldsymbol{\phi}_t^*) \notin \boldsymbol{c}_t$,we have $d_3(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t^*)) > 0$.

## C.17 Proof of Theorem 3.3.2

*Proof.* Let $\rho \in [0, 1]$ and $\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t') \in \boldsymbol{c}_t$. Recall that the set $\mathbb{m}(\boldsymbol{\phi}_t)$ contains all the SDF that prices the payoff over the infinitesimal time interval $[t, t + dt)$, and satisfies the asset pricing formula $\mathrm{E}[\boldsymbol{R}_t \boldsymbol{m}(\boldsymbol{\phi}_t)] = \boldsymbol{1}_A$, where $\boldsymbol{R}_t$ is the gross return vector from $t - dt$ to $t$. Then, the affine combination of $\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t')$ can be expressed in the form

$$\begin{aligned}\rho\, \mathbb{m}(\boldsymbol{\phi}_t) + (1 - \rho)\, \mathbb{m}(\boldsymbol{\phi}_t') &= \{\boldsymbol{m}(\boldsymbol{\phi}_t'') : \rho \mathrm{E}[\boldsymbol{R}_t \boldsymbol{m}(\boldsymbol{\phi}_t)] + (1 - \rho)\mathrm{E}[\boldsymbol{R}_1 \boldsymbol{m}(\boldsymbol{\phi}_t')] = \rho \boldsymbol{1}_A + (1 - \rho)\boldsymbol{1}_A\} \\ &= \{\boldsymbol{m}(\boldsymbol{\phi}_t'') : \rho \mathrm{E}[\boldsymbol{R}_t \boldsymbol{m}(\boldsymbol{\phi}_t)] + (1 - \rho)\mathrm{E}[\boldsymbol{R}_1 \boldsymbol{m}(\boldsymbol{\phi}_t')] = \boldsymbol{1}_A\} \in \boldsymbol{c}_t.\end{aligned}$$

Hence, $\boldsymbol{c}_t$ is *convex.*

Let $\mathbb{m}(\boldsymbol{\phi}_t) \in \boldsymbol{c}_t$ with $\boldsymbol{\phi}_t = (\boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx)) \in \boldsymbol{\Phi}_t$. There is $\eta > 0$ such that

$$\eta = \min_{\mathbb{m}(\boldsymbol{\phi}_t') \in \partial \boldsymbol{c}_t} \left\{ d_3 \left( \mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t') \right) \right\},$$

where $\boldsymbol{\phi}_t' \in \partial \boldsymbol{\phi}_t$. Then, since $\boldsymbol{c}_t = \bar{\boldsymbol{c}}_t \backslash \partial \boldsymbol{c}_t$, by choosing $\epsilon < \eta$, we have $\mathcal{B}(\mathbb{m}(\boldsymbol{\phi}_t), \epsilon) \subset \boldsymbol{c}_t$. Hence, $\boldsymbol{c}_t$ is *open* in $(\bar{\boldsymbol{c}}_t, d_3)$.

Next, we prove that $\boldsymbol{c}$ is bounded in $(\bar{\boldsymbol{c}}, d_3)$. Suppose there is a positive upper bound $R < \infty$ and let $(\mathbb{m}(\boldsymbol{\phi}_t^*), \mathbb{m}(\boldsymbol{\phi}_t'^*), w_t^*) = \arg\max_{\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}_t(\boldsymbol{\phi}_t') \in \boldsymbol{c}_t} d_3 \left( \mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}_t') \right)$. Then, any divergence of $(\boldsymbol{\phi}_t, \boldsymbol{\phi}_t')$ from $\boldsymbol{\phi}_t{}^*, (\boldsymbol{\phi}_t'^*)$ can be offset by the corresponding change in the optimal transport plan $w_t^*$, as $w_t^* \in W(P(\boldsymbol{\phi}_t), P(\boldsymbol{\phi}_t'))$ is a function of $(\boldsymbol{\phi}_t, \boldsymbol{\phi}_t')$ with

$$W(P(\boldsymbol{\phi}_t), P(\boldsymbol{\phi}_t')) := \left\{ w_t : \int w_t dP(\boldsymbol{\phi}_t') = P(\boldsymbol{\phi}_t), \int w_t dP(\boldsymbol{\phi}_t) = P(\boldsymbol{\phi}_t') \right\}.$$

Hence, $\boldsymbol{c}_t$ is *bounded.*

Lastly, to show that $\boldsymbol{c}_t$ is not compact, we just need one example of an open cover that has no finite open subcovers. Let $\{\boldsymbol{G}_{t,n}\} = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_{t,n}, n \in \mathbb{N}\}\}$, where

$$\{\boldsymbol{\Phi}_{t,n}\} = \left\{ \left( \boldsymbol{\mu}_t^B, \boldsymbol{\sigma}_t^B, \boldsymbol{\mu}_t^J, \boldsymbol{\sigma}_t^J, \boldsymbol{v}_t(dx) \right) : \boldsymbol{v}_t(dx) > \frac{1}{n}, n \in \mathbb{N} \right\}.$$

Thus, every element of $\boldsymbol{\phi}_t$ is in $\{\boldsymbol{\Phi}_{t,n}\}$ for some $n \in \mathbb{N}$, and therefore, every element of $\boldsymbol{c}_t$ is in $\{\boldsymbol{G}_{t,n}\}$ for some $n \in \mathbb{N}$, suggesting that $\boldsymbol{G}_t \subset \bigcap_{n=1}^{\infty} \{\boldsymbol{G}_{t,n}\}$. Moreover, since for any $n \in \mathbb{N}$, $\boldsymbol{G}_{t,n}$ has a neighbourhood centred at $\boldsymbol{G}_{t,n}$ with radius $\epsilon > 0$ lying in the set, $\{\boldsymbol{G}_{t,n}\}$ is an open cover of $\boldsymbol{c}_t$. Let $k, l \in \mathbb{N}$ such that $k > l > 2$, we have

$$\frac{1}{k} < \frac{1}{l} \Rightarrow \{\boldsymbol{\Phi}_{t,l}\} \subset \{\boldsymbol{\Phi}_{t,k}\} \Rightarrow \{\boldsymbol{G}_{t,l}\} \subset \{\boldsymbol{G}_{t,k}\}.$$

Hence, for any finite $m \in \mathbb{N}$, $\bigcup_{n=1}^{m} \{\boldsymbol{G}_{t,n}\} = \{\boldsymbol{G}_{t,m}\} = \{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \Phi_{t,m}, m \in \mathbb{N}\}$. However, for any $m \in \mathbb{N}$, there exists $\Phi_{t,m+1} \notin \{\boldsymbol{\Phi}_{t,m}\}$, while $\{\mathbb{m}(\boldsymbol{\phi}_t), \boldsymbol{\phi}_t \in \boldsymbol{\Phi}_{t,m+1}\} \in \boldsymbol{c}_t$. Therefore, $\{\boldsymbol{G}_{t,n}\}$ is an open cover of $\boldsymbol{c}_t$ that does not have a finite subcover. Hence, $\boldsymbol{c}_t$ is *not compact.* □

## C.18 Proof of Theorem 3.3.3

*Proof.* Let $t_0 \in (0,1]$, $\epsilon > 0$, we show that for any $t \in (0,1]$ such that $|t - t_0| < \delta$, we have

$$\begin{aligned}
&\left|MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t]}\right) - MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t_0]}\right)\right| \\
&= \left|\mathrm{E}_t\left[\min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right)\right] - \mathrm{E}_{t_0}\left[\min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right)\right]\right| \\
&= \left|\frac{1}{t}\int_0^t \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di - \frac{1}{t_0}\int_0^{t_0} \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di\right| \\
&\le \left|\frac{1}{t}\int_0^t \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di - \frac{1}{t}\int_0^{t_0} \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di\right| \qquad \text{(C.1)} \\
&+ \left|\frac{1}{t}\int_0^{t_0} \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di - \frac{1}{t_0}\int_0^{t_0} \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di\right| \\
&= \frac{1}{t}\left|\int_{t_0}^t \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di\right| + \left|\frac{1}{t} - \frac{1}{t_0}\right|\left|\int_0^{t_0} \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di\right| \\
&< \epsilon.
\end{aligned}$$

Thus, by choosing $\delta = \delta(t_0, \epsilon) > 0$ satisfying Equation (C.1), we have

$$\left|MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t]}\right) - MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t_0]}\right)\right| < \epsilon,$$

and therefore, $MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t]}\right)$ is *continuous* on the time interval $(0,1]$.

Next, we prove that $MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t]}\right)$ is not monotonic. Let

$$\begin{aligned}
F(t) &= MI\left(\{\boldsymbol{\phi}_i\}_{i\in(0,t]}\right) \\
&= \mathrm{E}_t\left[\min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right)\right] \\
&= \frac{1}{t}\int_0^t \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di,
\end{aligned}$$

$$\begin{aligned}
F'(t) &= -t^{-2}\int_0^t \min_{\boldsymbol{\phi}'_i\in\partial\boldsymbol{\Phi}_i} d_3\left(\mathbb{m}(\boldsymbol{\phi}_i), \mathbb{m}(\boldsymbol{\phi}'_i)\right) di + t^{-1}\min_{\boldsymbol{\phi}'_t\in\partial\boldsymbol{\Phi}_t} d_3\left(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}'_t)\right) \\
&= t^{-1}\left(\min_{\boldsymbol{\phi}'_t\in\partial\boldsymbol{\Phi}_t} d_3\left(\mathbb{m}(\boldsymbol{\phi}_t), \mathbb{m}(\boldsymbol{\phi}'_t)\right) - F(t)\right).
\end{aligned}$$

Therefore, whether the sign of $F'(t)$ depends on the difference between the sub-period market incompleteness at $t$ and the average of sub-periods market incompleteness up to $t$, which is not strictly increasing nor decreasing. Hence, $MI\left(\{\phi_i\}_{i\in(0,t]}\right)$ is *not monotonic.*

$\square$