Asset Allocation under the Basel Accord Risk Measures*

Zaiwen Wen† Xianhua Peng‡ Xin Liu§ Xiaodi Bai¶
Xiaoling Sun‖

First version January 2013, this version August 2013

Abstract

Financial institutions are currently required to meet more stringent capital requirements than they were before the recent financial crisis; in particular, the capital requirement for a large bank's trading book under the Basel 2.5 Accord more than doubles that under the Basel II Accord. The significant increase in capital requirements renders it necessary for banks to take into account the constraint of capital requirement when they make asset allocation decisions. In this paper, we propose a new asset allocation model that incorporates the regulatory capital requirements under both the Basel 2.5 Accord, which is currently in effect, and the Basel III Accord, which was recently proposed and is currently under discussion. We propose an unified algorithm based on the alternating direction augmented Lagrangian method to solve the model; we also establish the first-order optimality of the limit points of the sequence generated by the algorithm under some mild conditions. The algorithm is simple and easy to implement; each step of the algorithm consists of solving convex quadratic programming or

*We are grateful to Steven Kou for his insightful comments to the paper. Zaiwen Wen was partially supported by the NSFC grant 11101274 and research fund (20110073120069) for the Doctoral Program of Higher Education of China. Xianhua Peng was partially supported by a grant from School-Based-Initiatives of HKUST (Grant No. SBI11SC03) and Hong Kong RGC Direct Allocation Grant (Project No. DAG12SC05-3). Xin Liu was partially supported by the NSFC grants 10831006 and 11101409. Xiaodi Bai and Xiaoling Sun were supported by the NSFC grant 10971034 and the Joint NSFC/RGC grant 71061160506.

†Department of Mathematics, MOE-LSC and Institute of Natural Sciences, Shanghai Jiaotong University, Shanghai 200240, China, zw2109@sjtu.edu.cn.

‡Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, maxhpeng@ust.hk.

§Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China, liuxin@lsec.cc.ac.cn.

¶Department of Management Science, School of Management, Fudan University, Shanghai 200433, China, xdbai@fudan.edu.cn.

‖Department of Management Science, School of Management, Fudan University, Shanghai 200433, China, xls@fudan.edu.cn.

one-dimensional subproblems. Numerical experiments on simulated and real market data show that the algorithm compares favorably with other existing methods, especially in cases in which the model is non-convex.

Keywords: Asset Allocation, Basel Accords, Capital Requirements, Value-at-Risk, Conditional Value-at-Risk, Expected Shortfall, Alternating Direction Augmented Lagrangian Methods

1 Introduction

One of the major consequences of the financial crisis that began in 2007 is that financial institutions are now required to meet more stringent capital requirements than they were before the crisis. The considerable increase in capital requirements has been imposed through the Basel Accords, which have undergone substantial revision since the inception of the financial crisis. The framework of the latest version of the Basel Accord, the Basel III Accord (Basel Committee on Banking Supervision, 2010), was announced in December 2010 and is soon to be implemented in many leading nations, including the United States (Board of Governors of the Federal Reserve Systems, 2012).

In particular, the capital requirements for banks' trading books, which are calculated by the Basel Accord risk measure for the trading book, have been increased substantially. Before the 2007 financial crisis, the Basel II risk measure (Basel Committee on Banking Supervision, 2006) was used in the calculation. During the crisis, it was found that the Basel II risk measure had serious drawbacks, such as being procyclical and not being conservative enough. In response to the financial crisis, the Basel committee revised the Basel II market risk framework and imposed the "Basel 2.5" risk measure (Basel Committee on Banking Supervision, 2009) in July 2009. It has been estimated that the capital requirement for a large bank's trading book under the Basel 2.5 risk measure on average *more than doubles* that under the Basel II risk measure (Basel Committee on Banking Supervision, 2012, p. 11).

The substantial increase in the capital requirements for the trading book makes it more important for banks to take into account the constraint of capital requirements when they construct investment portfolios. In this paper, we address this issue by proposing a new asset allocation model that incorporates the capital requirement imposed by the Basal Accords. More precisely, we propose the *"mean-ρ-Basel"* asset allocation model, in which ρ denotes the risk measure used for measuring the risk of the investment portfolio, such as variance, value-at-risk (VaR), or conditional value-at-risk (CVaR); ρ can be freely chosen by the portfolio manager; and "Basel" denotes the constraint that the regulatory capital of the portfolio calculated by the Basel Accord risk measure should not exceed a certain upper limit.

The complexity of the Basel Accord risk measures for calculating the capital requirements poses a challenge to solving the proposed "mean-ρ-Basel" asset allocation model. The Basel Accords use VaR or CVaR with *scenario analysis* as the risk measure to calculate the capital requirements for a bank's trading book. Scenario analysis is used to analyze the behavior of random losses under different scenarios; a scenario refers to a specific economic regime such as an economic boom and a financial crisis. The Basel II risk measure involves the calculation of VaR under 60 different scenarios. The Basel 2.5 risk measure involves the calculation of VaR under 120 scenarios, including 60 stressed scenarios. Most recently, in May 2012, the Basel Committee released a consultative document (Basel Committee on Banking Supervision, 2012) that presents the initial policy proposal of a new risk measure to replace the Basel 2.5 risk measure for the trading book; the new risk measure involves the calculation of CVaR under stressed scenarios. Currently, this new proposal is under discussion and has not been finalized. It is beyond the scope of this paper to discuss whether the newly proposed risk measure is superior to the Basel 2.5 risk measure; hence, we will consider both the Basel 2.5 and the newly proposed Basel risk measure in the mean-ρ-Basel model. See Section 2.2 for details regarding the Basel Accord risk measures.

Numerous studies have examined the single-period asset allocation model of *"mean-ρ"*, in which ρ is a measure of portfolio risk such as variance, VaR, or CVaR. On recent developments in the mean-variance asset allocation models and associated algorithms, see e.g., Chairawongse et al. (2012). Iyengar and Ma (2013) propose a fast iterative gradient descent algorithm capable of handling large-scale problems for the mean-CVaR problem. Lim, Shanthikumar, and Vahn (2011) evaluate CVaR as the risk measure in data-driven portfolio optimization and show that portfolios obtained by solving mean-CVaR problems are unreliable due to estimation errors of CVaR and/or the mean asset returns. To address the issue of estimation risk, Karoui, Lim, and Vahn (2011) introduce a new approach, called performance-based regularization, to the data-driven mean-CVaR portfolio optimization problem. Rockafellar and Uryasev (2002) develop a method to reduce the data-driven mean-CVaR asset allocation problem to a linear programming (LP) problem. The mean-VaR problem is more difficult than the mean-CVaR due to the non-convexity of VaR. Software packages such as CPLEX can be used to solve small-to-medium sized problems of this type. Recently, Cui et al. (2013) propose a second-order cone programming method to solve a mean-VaR model when VaR is estimated by its first-order or second-order approximations. Bai et al. (2012) propose a penalty decomposition method for probabilistically constrained programs including the mean-VaR problem.

It appears to be more challenging to solve the mean-ρ-Basel model than the mean-ρ model due to the complexity of the Basel Accord risk measures that involve multiple VaRs or CVaRs under various scenarios and the non-convexity of VaR. In this paper, we develop an unified and computationally efficient method to solve the mean-ρ-Basel problem. This method is based on the *alternating direction augmented Lagrangian method* (ADM) (see, e.g., Wen, Goldfarb, and Yin 2010; He and Yuan 2012; Hong and Luo 2012 and the references therein). The method is very simple and easy to implement; it reduces the original problem to one-dimensional optimization or convex quadratic programming subproblems that may even have closed-form solutions; hence, the method is capable of solving large scale problems. When the mean-ρ-Basel problem is convex for some specific ρ and Basel constraint (e.g., ρ is variance and the Basel constraint is specified by the newly proposed Basel Accord risk measure), the method is guaranteed to converge to the globally optimal solution; when the problem is non-convex, we show that the limit points of the sequence generated by the method satisfy the first-order optimality condition. See Section 4.

The proposed method also applies to mean-ρ problems such as the mean-VaR, mean-CVaR, and "*mean-Basel*" problem, in which the Basel Accord risk measures are used to quantify the risk of the portfolio. The Basel Accord risk measures involve multiple VaR or CVaR under different scenarios, which essentially correspond to different models or distributions of asset returns. Hence, using the Basel Accord risk measures, or, more generally, VaR or CVaR with scenario analysis, as the portfolio risk measure provides a way to address the problem of model uncertainty.

In summary, the main contribution of the paper is two-fold. (i) We formulate a new asset allocation model, the mean-ρ-Basel model, which takes into account the regulatory capital constraint specified by the Basel Accord risk measure for trading books. We also formulate and study the related mean-Basel model. To the best of our knowledge, there has been no literature on asset allocation involving the Basel Accord risk measures. (ii) We propose an efficient alternating direction augmented Lagrangian method for solving the mean-ρ-Basel and mean-ρ models. For non-convex cases of these models, we establish the first-order optimality of the limit points of iterative sequence generated by the method under mild conditions. Although there is no theoretical guarantee that the method will converge to the global solution in non-convex cases of these models, numerical experiments on simulated and real market data show that the method can identify suboptimal solutions that can often be superior to the approximate solutions of the mixed-integer programming formulation computed by CPLEX within one hour.

The remainder of the paper is organized as follows. In Section 2, we review the definition and properties of the Basel Accord risk measures for trading books as well as some other relevant risk measures. In Section 3, we formulate the mean-ρ-Basel asset allocation model in which the Basel Accord risk measures are used for setting a regulatory capital constraint. In Section 4, we propose the alternating direction augmented Lagrangian method for solving the mean-ρ-Basel and the mean-ρ problems; we also provide convergence analysis of the method. Section 5 provides the numerical results, which demonstrate the accuracy and efficiency of the proposed method.

2 Review of Relevant Risk Measures

Variance is probably the best-known risk measure; in addition to variance, there is a vast literature on theoretical frameworks and concrete examples of risk measures. As it is beyond the scope of this paper to discuss and compare different risk measures, we review only the risk measures that are used in the asset allocation problems considered in this paper.

2.1 Value-at-Risk and Conditional Value-at-Risk (Expected Shortfall)

Value-at-Risk (VaR) is one of the most widely used risk measures in risk management. VaR is a quantile of the loss distribution at some pre-defined probability level. More precisely, let $F_X(\cdot)$ be the distribution function of the random loss X, then, for a given $\alpha \in (0, 1)$, VaR of X at level α is defined as

$$\text{VaR}_\alpha(X) := \inf\{x \mid F_X(x) \geq \alpha\} = F_X^{-1}(\alpha). \tag{1}$$

Jorion (2007) provides a comprehensive discussion of VaR and risk management.

Conditional Value-at-Risk (CVaR), proposed by Rockafellar and Uryasev (2002), is another prominent and widely used risk measure. For the random loss X, the α-tail distribution function of X is defined as

$$F_{\alpha,X}(x) := \begin{cases} 0, & \text{for } x < \text{VaR}_\alpha(X), \\ \frac{F_X(x)-\alpha}{1-\alpha}, & \text{for } x \geq \text{VaR}_\alpha(X). \end{cases} \tag{2}$$

Then, the CVaR at level α of X is defined as

$$\text{CVaR}_\alpha(X) := \text{mean of the } \alpha\text{-tail distribution of } X = \int_{-\infty}^{\infty} x dF_{\alpha,X}(x). \tag{3}$$

Expected shortfall (ES) is a risk measure that is equivalent to CVaR and that was introduced independently in Acerbi and Tasche (2002). CVaR and ES have the subadditivity property and belong to the class of coherent risk measures (Artzner et al., 1999); VaR may not satisfy subadditivity and belongs to another class of risk measures called insurance risk measures (Wang, Young, and Panjer, 1997).

2.2 Basel Accord Risk Measures for Trading Books

The Basel Accords use VaR or CVaR with scenario analysis as the risk measure for calculating capital requirements for banks' trading books. A scenario refers to a specific economic regime, such as an economic boom or a financial crisis. Scenario analysis is necessary because studies have shown that the behavior of economic variables is substantially different under different economic regimes (see, e.g., Hamilton, 1989). In particular, many economic variables exhibit dramatic changes in their behavior during financial crises (Hamilton, 2005) or when government monetary or fiscal policies undergo sudden changes (Sims and Zha, 2006). There is also evidence that the volatility and correlation among asset returns increase in economic downturns (see, e.g., Dai, Singleton, and Yang, 2007).

The Basel II Accord (Basel Committee on Banking Supervision, 2006) specifies that the capital charge for the trading book on any particular day t for banks using the internal models approach should be calculated by the formula

$$c_t = \max\left\{\text{VaR}_{\alpha,t-1}(X), \frac{k}{60}\sum_{s=1}^{60}\text{VaR}_{\alpha,t-s}(X)\right\}, \tag{4}$$

where X is the loss of the bank's trading book; k is a constant that is no less than 3; $\text{VaR}_{\alpha,t-s}(X)$ is the 10-day VaR of X at $\alpha = 99\%$ confidence level calculated on day $t-s$, $s = 1, \ldots, 60$. $\text{VaR}_{\alpha,t-s}(X)$ is calculated under the scenario corresponding to information available on day $t-s$. For example, $\text{VaR}_{\alpha,t-s}(X)$ of a portfolio of equity options is calculated conditional on the value of the equity prices, equity volatilities, yield curves, etc., on day $t-s$. Therefore, the Basel II risk measure is a VaR with scenario analysis that involves 60 scenarios.

Since the 2007 financial crisis, the Basel II risk measure (4) has been criticized for two reasons: (i) This risk measure is based on contemporaneous observations and hence is *procyclical*, i.e., risk measurement obtained by it tend to be low in booms and high in crises, which is exactly opposite to the goal of effective regulation (Adrian and Brunnermeier, 2008). (ii) This risk measure is not conservative enough. In fact, banks' actual losses during the financial crisis were significantly higher than the capital requirements calculated by the risk measure.

In response to the financial crisis, the Basel Committee revised the Basel II market risk framework and replaced the Basel II risk measure with the "Basel 2.5" risk measure in July 2009 (Basel Committee on Banking Supervision, 2009). The Basel 2.5 risk measure for calculating capital requirements for trading books is defined by

$$
\begin{aligned}
c_t = & \max\left\{\text{VaR}_{\alpha,t-1}(X), \frac{k}{60}\sum_{s=1}^{60}\text{VaR}_{\alpha,t-s}(X)\right\} \\
& + \max\left\{\text{sVaR}_{\alpha,t-1}(X), \frac{\ell}{60}\sum_{s=1}^{60}\text{sVaR}_{\alpha,t-s}(X)\right\}, \qquad (5)
\end{aligned}
$$

where $\text{VaR}_{\alpha,t-s}(X)$ is the same as that in (4); k and ℓ are constants no less than 3; and $\text{sVaR}_{\alpha,t-s}(X)$ is called the *stressed* VaR of X on day $t-s$ at confidence level $\alpha = 99\%$, which is calculated under a scenario in which the financial market is under significant stress, such as the one that happened during the period from 2007 to 2008. The additional capital requirements based on stressed VaR help to reduce the procyclicality of the Basel II risk measure (4) and significantly increase the capital requirements.

In May 2012, the Basel Committee released a consultative document (Basel Committee on Banking Supervision, 2012) that presents the initial policy proposal regarding the Basel Committee's fundamental review of the trading book capital requirements. In particular, the Committee proposed a new risk measure to replace the Basel 2.5 risk measure; the new risk measure uses CVaR (or, equivalently, ES) instead of VaR to calculate capital requirements. More precisely, under the new risk measure, the capital requirement for a group of trading desks that share similar major risk factors, such as equity, credit, interest rate, and currency, is defined as the CVaR of the loss that may be incurred by the group of trading desks; the CVaR should be calculated under stressed scenarios rather than under current market conditions. For example, an equity trading desk and an equity option trading desk would be grouped together for the purpose of calculating regulatory capital. This proposed risk measure is currently under discussion, and it is not yet clear whether it is going to be the final version of the Basel III risk measure. In addition, the proposal has not clearly stated if the capital charge for the tth day will depend solely on the stressed CVaR calculated on day $t-1$ or on the CVaR calculated on day $t-s$ for $s = 1, 2, \ldots, 60$, as in Basel 2.5. To be more consistent with Basel 2.5, we consider the following "Basel III" risk measure:

$$c_t = \max\left\{ \text{sCVaR}_{\alpha,t-1}, \frac{\ell}{60}\sum_{s=1}^{60} \text{sCVaR}_{\alpha,t-s} \right\}, \tag{6}$$

where $\text{sCVaR}_{\alpha,t-s}$ is the *stressed* CVaR at level α calculated on day $t-s$. The proposal suggests specifying α to be a level smaller than 99% due to the difficulty of estimating CVaR at high confidence levels, but the exact value of α has not been determined. In the numerical examples of Section 5, we choose $\alpha = 98\%$.

3 A New Asset Allocation Model Incorporating the Basel Accord Capital Constraint

Consider a portfolio composed of d assets and let $u = (u_1, u_2, \ldots, u_d)^\top \in \mathbb{R}^d$ denote the portfolio weights of these assets, which are the percentage of initial wealth invested in the assets. Let $R = (R_1, R_2, \ldots, R_d)^\top \in \mathbb{R}^d$ be the random vector of simple returns of these assets over a specified time horizon, e.g., one day. Then the simple return of the portfolio is $R^\top u$ and $-R^\top u$ is the loss of the portfolio (per \$1 of investment). Let $\mu \in \mathbb{R}^d$ be the (estimated) expected returns of the d assets. Then $\mu^\top u$ is the expected return of the portfolio.

The risk of the portfolio is measured by $\rho(-R^\top u)$, where ρ is a properly chosen risk measure. There are generally two approaches to the computation of $\rho(-R^\top u)$: (i) one first assumes and estimates a (parametric) probability model for the joint distribution of R and then computes $\rho(-R^\top u)$; (ii) one estimates the risk $\rho(-R^\top u)$ directly from the historical observations of R without assuming any hypothetical model for R.

As discussed in Section 2.2, the return vector R is usually observed under different scenarios, such as economic booms and financial crises. Suppose there are m scenarios. For each $s = 1, \ldots, m$, let $\tilde{R}^{[s]} \in \mathbb{R}^{n_s \times d}$ be the collection of n_s observations of R under the sth scenario, where each row of $\tilde{R}^{[s]}$ represents one observation of $R^\top$. Then, we define the matrix $\tilde{R}$ and the observations of portfolio loss $x(u)$ as follows:

$$\tilde{R} := \begin{pmatrix} \tilde{R}^{[1]} \\ \tilde{R}^{[2]} \\ \vdots \\ \tilde{R}^{[m]} \end{pmatrix} \in \mathbb{R}^{n\times d},\ x(u) := -\tilde{R}u = \begin{pmatrix} -\tilde{R}^{[1]}u \\ -\tilde{R}^{[2]}u \\ \vdots \\ -\tilde{R}^{[m]}u \end{pmatrix} = \begin{pmatrix} x^{[1]}(u) \\ x^{[2]}(u) \\ \vdots \\ x^{[m]}(u) \end{pmatrix} \in \mathbb{R}^n,\ n := \sum_{s=1}^{m} n_s, \tag{7}$$

where $x^{[s]}(u) := -\tilde{R}^{[s]}u \in \mathbb{R}^{n_s}$ denotes the observations of portfolio loss under the sth scenario, $s = 1, 2, \ldots, m$.

In this paper, we estimate $\rho(-R^\top u)$ directly from the return observations $\tilde{R}$, as this approach does not require a subjective model for R and hence greatly reduces model misspecification error.

3.1 Sample Versions of Measures of Portfolio Risk

In the following, we use $\lceil \cdot \rceil$ to denote the ceiling function. For $x = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^n$, let $(i_1, i_2, \ldots, i_n)$ be a permutation of $(1, 2, \ldots, n)$ such that $x_{i_1} \le x_{i_2} \le \cdots \le x_{i_n}$. Then, we define $x_{(j)} := x_{i_j}$, $j = 1, \ldots, n$; hence, $x_{(j)}$ denotes the jth smallest component of x.

Given the observation $\tilde{R}^{[s]}$, the empirical distribution function of $(-R^\top u)$ under scenario s is given by

$$\hat{F}^{[s]}_{(-R^\top u)}(y) := \frac{1}{n_s} \sum_{i=1}^{n_s} 1_{\{x^{[s]}(u)_i \le y\}}. \tag{8}$$

Then, for each risk measure ρ discussed in Section 2.2, $\rho(-R^\top u)$ can be estimated from the return observations $\tilde{R}$ by substituting $\hat{F}^{[s]}_{(-R^\top u)}(\cdot)$ for the distribution function of $(-R^\top u)$ under each scenario s. Thus, we obtain the following sample versions of risk measures.

Variance: Suppose there is one scenario, i.e., $m = 1$. Then the sample variance of portfolio return is

$$\rho_{\text{Variance}}(x(u)) = \frac{1}{n} x(u)^\top x(u) - \frac{1}{n^2} x(u)^\top \mathbf{1}\mathbf{1}^\top x(u), \text{ where } \mathbf{1} := (1, 1, \ldots, 1)^\top \in \mathbb{R}^n. \tag{9}$$

VaR: Suppose that $m = 1$. For a given $\alpha \in (0, 1)$, let $p = \lceil \alpha n \rceil$. Then the sample VaR at level α of the portfolio is

$$\rho_{\text{VaR}_\alpha}(x(u)) := x(u)_{(p)} = (-\tilde{R}u)_{(p)}. \tag{10}$$

CVaR: Suppose that $m = 1$. For a given $\alpha \in (0, 1)$, let $p = \lceil \alpha n \rceil$. Then the sample CVaR at level α of the portfolio is

$$\rho_{\text{CVaR}_\alpha}(x(u)) := \frac{p - \alpha n}{(1-\alpha)n} x(u)_{(p)} + \frac{1}{(1-\alpha)n} \sum_{i=p+1}^{n} x(u)_{(i)}. \tag{11}$$

By Theorem 10 in Rockafellar and Uryasev (2002), $\rho_{\text{CVaR}_\alpha}(x(u))$ can also be represented by

$$\rho_{\text{CVaR}_\alpha}(x(u)) = \min_{t \in \mathbb{R}} \; t + \frac{1}{(1-\alpha)n} \sum_{i=1}^{n} (x(u)_i - t)_+, \text{ where } y_+ := \max(y, 0). \tag{12}$$

Basel 2.5: For a given $\alpha \in (0,1)$, let $p_s = \lceil \alpha n_s \rceil$, $s = 1, \ldots, m$. Then $x^{[s]}(u)_{(p_s)}$ is the sample VaR at level α of the portfolio estimated from the data set $\tilde{R}^{[s]}$. Let $m_1 = m_2 = 60$ and $m = 120$. Suppose the first m_1 scenarios correspond to current market conditions and the last m_2 scenarios correspond to stressed scenarios. Then, the sample version of the Basel 2.5 Accord risk measure is given by

$$\rho_{\text{Basel2.5}}(x(u)) := \max\left\{x^{[1]}(u)_{(p_1)}, \frac{k}{m_1}\sum_{s=1}^{m_1} x^{[s]}(u)_{(p_s)}\right\} + \max\left\{x^{[m_1+1]}(u)_{(p_{m_1+1})}, \frac{\ell}{m_2}\sum_{s=m_1+1}^{m} x^{[s]}(u)_{(p_s)}\right\}. \tag{13}$$

Basel III: Let α and p_s be defined previously. Then

$$\rho_{\text{CVaR}_\alpha}(x^{[s]}(u)) := \frac{p_s - \alpha n_s}{(1-\alpha)n_s} x^{[s]}(u)_{(p_s)} + \frac{1}{(1-\alpha)n_s}\sum_{i=p_s+1}^{n_s} x^{[s]}(u)_{(i)} \tag{14}$$

is the sample CVaR at level α of the portfolio estimated from the data set $\tilde{R}^{[s]}$. Suppose the first $m_1 = 60$ scenarios correspond to current market conditions and the last $m_2 = 60$ scenarios correspond to stressed scenarios. Then the sample version of the Basel-III risk measure is

$$\rho_{\text{Basel3}}(x(u)) := \max\left\{\rho_{\text{CVaR}_\alpha}(x^{[m_1+1]}(u)), \frac{\ell}{m_2}\sum_{s=m_1+1}^{m} \rho_{\text{CVaR}_\alpha}(x^{[s]}(u))\right\}. \tag{15}$$

3.2 The "Mean-ρ-Basel" Asset Allocation Model

Suppose a portfolio manager in a financial institution attempts to construct a portfolio composed of the d assets and to choose the portfolio weights $u \in \mathbb{R}^d$ to optimize the portfolio performance. The manager can freely choose a risk measure ρ to measure the risk of the portfolio, such as variance, VaR, or CVaR; in addition, he or she has the freedom to choose a model for the asset returns R or a data set $\tilde{Y} \in \mathbb{R}^{n' \times d}$, which has a similar structure to that of $\tilde{R}$ defined in (7) and contains observations of the asset returns, to estimate the portfolio risk. Hence, the portfolio risk will be given by $\rho(y(u))$, where $y(u) := -\tilde{Y}u$. Furthermore, the manager can specify that the expected portfolio return should be no less than a target return r_0, namely, the portfolio weights u should satisfy

$$u \in \mathcal{U}_{r_0} := \{u \in \mathbb{R}^d \mid \mu^\top u \geq r_0, \mathbf{1}^\top u = 1, u \geq 0\}.$$

Here, it is assumed that the portfolio is long only; this assumption can be relaxed or removed without incurring additional technical difficulty in solving the asset allocation problem specified below.

At the same time, the manager has to meet the constraint that the regulatory capital for his or her portfolio should not exceed an upper limit C_0, which is allocated to him or her by the financial institution's senior management. The capital requirement for the portfolio is calculated by the Basel Accord risk measure ρ_{Basel}, which is specified by the regulators; in addition, the data set $\tilde{R}$ used for calculating the capital requirements should also satisfy certain criteria and cannot be freely chosen by the portfolio manager. For example, the Basel 2.5 risk measure requires that $\tilde{R}$ should include 60 normal scenarios and 60 stressed scenarios. Hence, the data set $\tilde{R}$ may be different from the data set $\tilde{Y}$, and the capital requirement for the portfolio is $\rho_{\text{Basel}}(x(u))$, where $x(u) = -\tilde{R}u$.

To address the concerns of the portfolio manager, we propose the following ***"mean-ρ-Basel"*** asset allocation problem:

$$\begin{aligned} &\min_{u \in \mathcal{U}_{r_0}} \ \rho(y(u)) \\ &\text{s.t.} \quad \rho_{\text{Basel}}(x(u)) \leq C_0, \end{aligned} \tag{16}$$

where $x(u) = -\tilde{R}u$; $y(u) = -\tilde{Y}u$; ρ_{Basel} is the Basel Accord risk measure for calculating regulatory capital, i.e., $\rho_{\text{Basel2.5}}$ or ρ_{Basel3}; C_0 is the upper bound of the available capital; and ρ is the risk measure that the manager chooses for gauging the risk of the portfolio, such as variance, VaR, or CVaR.

The mean-ρ-Basel problem (16) with $\rho = \rho_{\text{VaR}_\alpha}$ or $\rho_{\text{Basel}} = \rho_{\text{Basel2.5}}$ is non-convex and is usually difficult to solve, as it can be formulated as a mixed-integer programming (MIP) problem. For example, by introducing $z' \in \{0,1\}^{n'}$ and $z^{[s]} \in \{0,1\}^{n_s}$ for $1 \leq s \leq m$, the mean-VaR-Basel2.5 problem can be formulated as the following MIP problem:

$$\begin{aligned} \min_{u,z,\beta,\gamma} \quad & \beta_0 \\ \text{s.t.} \quad & -\tilde{Y}u \leq \beta_0 \mathbf{1} + \eta z', \mathbf{1}^\top z' \leq n' - p', \ z' \in \{0,1\}^{n'}, \\ & -\tilde{R}^{[s]}u \leq \beta_s \mathbf{1} + \eta z^{[s]}, \mathbf{1}^\top z^{[s]} \leq n_s - p_s, z^{[s]} \in \{0,1\}^{n_s}, \ s = 1, \ldots, m, \\ & \beta_1 \leq \gamma_1, \ \frac{k}{m_1} \sum_{s=1}^{m_1} \beta_s \leq \gamma_1, \beta_{m_1+1} \leq \gamma_2, \ \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \beta_s \leq \gamma_2, \\ & \gamma_1 + \gamma_2 \leq C_0, \\ & u \in \mathcal{U}_{r_0}, \end{aligned} \tag{17}$$

where $p' := \lceil \alpha n' \rceil$, $p_s := \lceil \alpha n_s \rceil$, η is a large constant. For instance, η can be chosen to be $\eta = \max_{u \in \mathcal{U}_{r_0}} \max_{j=1,\ldots,n} (-\tilde{Y}u)_j$.

Similarly, by (12), the mean-VaR-Basel3 problem can be formulated as the following

MIP problem:

$$\begin{aligned}
\min_{u,z',\beta_0,t,r} \quad & \beta_0 \\
\text{s.t.} \quad & -\tilde{Y}u \le \beta_0 \mathbf{1} + \eta z', \mathbf{1}^\top z' \le n' - p', \; z' \in \{0,1\}^{n'}, \\
& t_{m_1+1} + \frac{1}{(1-\alpha_3)n_{m_1+1}} \sum_{i=1}^{n_{m_1+1}} r_i^{[m_1+1]} \le C_0, \\
& \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} (t_s + \frac{1}{(1-\alpha_3)n_s} \sum_{i=1}^{n_s} r_i^{[s]}) \le C_0, \\
& r_i^{[s]} \ge 0, \; r_i^{[s]} \ge -\tilde{R}_i^{[s]} u - t_s, i = 1, \ldots, n_s, s = m_1+1, \ldots, m, \\
& u \in \mathcal{U}_{r_0}.
\end{aligned} \tag{18}$$

On the other hand, the mean-ρ-Basel problem with ρ_{Basel} being ρ_{Basel3} and with ρ being ρ_{Variance} or $\rho_{\text{CVaR}_\alpha}$ is convex. More precisely, the mean-variance-Basel3 and mean-CVaR-Basel3 problems can be formulated as a quadratic programming (QP) problem and a linear programming (LP) problem, respectively, thanks to the LP formulation of CVaR given in (12).

We develop a unified method for solving the mean-ρ-Basel problem in Section 4.1 and provide convergence analysis of the method in Section 4.2. The method can also be applied to solve the classical *"mean-ρ"* problem:

$$\min_{u \in \mathcal{U}_{r_0}} \; \rho(x(u)), \tag{19}$$

where ρ can be any risk measure chosen by the portfolio manager, such as variance, VaR, CVaR, and ρ_{Basel}. If $\rho = \rho_{\text{Basel}}$, problem (19) is the *"mean-Basel"* problem, in which the Basel Accord risk measures are used to quantify the risk of the portfolio. The Basel Accord risk measures involve multiple VaR or CVaR under different scenarios, which essentially correspond to different models or distributions of asset returns. Hence, using the Basel Accord risk measures, or, more generally, VaR or CVaR with scenario analysis, as the portfolio risk measure provides a way to address the problem of model uncertainty. Alternatively, the portfolio manager can construct the portfolio by maximizing the expected return of portfolio subject to the constraint that the portfolio risk, measured by ρ, does not exceed a pre-specified risk budget b_0. The corresponding asset allocation problem is

$$\begin{aligned}
\min_{u \in \mathcal{U}} \; & -\mu^\top u \\
\text{s.t.} \; & \rho(x(u)) \le b_0,
\end{aligned} \tag{20}$$

where $\mathcal{U} = \{u \in \mathbb{R}^d \mid \mathbf{1}^\top u = 1, u \ge 0\}$. The mean-$\rho$ problems (19) and (20) with $\rho \in \{\rho_{\text{VaR}_\alpha}, \rho_{\text{Basel2.5}}\}$ are also MIP problems which are difficult to solve. The details of the method for solving these problems are given in Section 4.3.

4 The Alternating Direction Augmented Lagrangian Method

In this section, we propose a unified algorithm adapted from the *alternating direction augmented Lagrangian* method (ADM) to solve the mean-ρ-Basel and the mean-ρ problem. Although the ADM approach has been used in convex optimization (see, e.g., Wen, Goldfarb, and Yin 2010; He and Yuan 2012; and Hong and Luo 2012), it appears that its use for solving non-convex problems involving VaR or Basel Accords risk measures is new. In particular, the proposed method is different from the penalty decomposition methods proposed in Bai et al. (2012), in which the division of blocks of variables leads to subproblems that are more expensive to solve.

4.1 The ADM Algorithm for Solving the Mean-ρ-Basel Problem (16)

The problem (16) is equivalent to

$$\begin{aligned} \min_{u\in\mathcal{U}_{r_0},x\in\mathbb{R}^n,y\in\mathbb{R}^{n'}} \quad & \rho(y) \\ \text{s.t.} \quad & \rho_{\text{Basel}}(x) \le C_0, \\ & x + \tilde{R}u = 0, \\ & y + \tilde{Y}u = 0. \end{aligned} \tag{21}$$

We then define the augmented Lagrangian function for (21) as follows:

$$\mathcal{L}(x,y,u,\lambda,\pi) := \rho(y) + \lambda^\top(x+\tilde{R}u) + \frac{\sigma_1}{2}\|x+\tilde{R}u\|^2 + \pi^\top(y+\tilde{Y}u) + \frac{\sigma_2}{2}\|y+\tilde{Y}u\|^2, \tag{22}$$

where $\sigma_1, \sigma_2 > 0$ is the penalty parameter and $\lambda \in \mathbb{R}^n$ and $\pi \in \mathbb{R}^{n'}$ are the Lagrangian multipliers associated with the equality constraints $x + \tilde{R}u = 0$ and $y + \tilde{Y}u = 0$, respectively.

We propose an ADM algorithm that minimizes (22) with respect to x, y, and u in an alternating fashion while updating λ and π in the iteration. More precisely, let $x^{(j)}$, $y^{(j)}$, and $u^{(j)}$ be the values of x, y, and u at the beginning of the jth iteration of the algorithm; then the algorithm updates the values of x, y, and u by solving the following three subproblems sequentially:

$$x^{(j+1)} = \arg\min_{x\in\mathbb{R}^n} \ \mathcal{L}(x, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)}), \ \text{s.t.}\ \rho_{\text{Basel}}(x) \le C_0, \tag{23}$$

$$y^{(j+1)} = \arg\min_{y\in\mathbb{R}^{n'}} \ \mathcal{L}(x^{(j+1)}, y, u^{(j)}, \lambda^{(j)}, \pi^{(j)}), \tag{24}$$

$$u^{(j+1)} = \arg\min_{u\in\mathcal{U}_{r_0}} \ \mathcal{L}(x^{(j+1)}, y^{(j+1)}, u, \lambda^{(j)}, \pi^{(j)}). \tag{25}$$

Then, it updates the the Lagrangian multipliers by

$$\lambda^{(j+1)} = \lambda^{(j)} + \beta_1\sigma_1(x^{(j+1)} + \tilde{R}u^{(j+1)}), \tag{26}$$

$$\pi^{(j+1)} = \pi^{(j)} + \beta_2\sigma_2(y^{(j+1)} + \tilde{Y}u^{(j+1)}), \tag{27}$$

where $\beta_1, \beta_2 > 0$ are appropriately chosen step lengths.

The solutions to problems (23) and (24) are given in the lemmas at the end of this subsection; these solutions are obtained either in closed form, by solving QP problems, or by minimizing a single variable function on a closed interval. As for problem (25), simple algebra shows that it is equivalent to the following QP problem (28):

$$\begin{aligned} u^{(j+1)} &= \arg\min_{u\in\mathcal{U}_{r_0}} \frac{1}{2}u^\top(\sigma_1\tilde{R}^\top\tilde{R} + \sigma_2\tilde{Y}^\top\tilde{Y})u + b_e^\top u, \text{ where} \\ b_e &= \tilde{R}^\top(\lambda^{(j)} + \sigma_1 x^{(j+1)}) + \tilde{Y}^\top(\pi^{(j)} + \sigma_2 y^{(j+1)}). \end{aligned} \tag{28}$$

The complete ADM algorithm is given as follows.

Algorithm 1 ADM algorithm for solving the mean-ρ-Basel problem (16)

1: Choose parameter $\sigma_1 > 0$, $\sigma_2 > 0$, $\beta_1 > 0$, $\beta_2 > 0$
2: Set $j = 0$; initialize $y^{(0)} \in \mathbb{R}^{n'}$, $u^{(0)} \in \mathbb{R}^d$, $\lambda^{(0)} := 0$, and $\pi^{(0)} := 0$
3: **while** $\{u^{(j)}\}$ has not converged **do**
4: update $x^{(j+1)}$ to be the solution to problem (23); the solution is given in Lemma 4.1 and Lemma 4.2 for $\rho_{\text{Basel}} = \rho_{\text{Basel2.5}}$ and $\rho_{\text{Basel}} = \rho_{\text{Basel3}}$, respectively
5: update $y^{(j+1)}$ to be the solution to problem (24); the solution is given in Lemma 4.3, Lemma 4.4, and Lemma 4.5 for $\rho = \rho_{\text{Variance}}$, $\rho = \rho_{\text{VaR}_\alpha}$, and $\rho = \rho_{\text{CVaR}_\alpha}$, respectively
6: update $u^{(j+1)}$ by solving the QP problem (28)
7: update $\lambda^{(j+1)}$ and $\pi^{(j+1)}$ by (26) and (27), respectively
8: increase j by one and continue.
9: **end while**

The algorithm is very simple and easy to implement. Standard QP solvers, such as CPLEX, can be used to solve the QP problems in Step 4, 5, and 6 of the algorithm. In step 5 for the case of $\rho = \rho_{\text{CVaR}_\alpha}$, the solution is obtained by minimizing a single-variable function on a closed interval, which can be solved by golden section search and parabolic interpolation (e.g., the function "fminbnd" in Matlab).

One particular implementation of the ADM algorithm including the specification of the parameters σ_i and β_i and the convergence test is given in Section 5.2.

The lemmas for solving the subproblems in the algorithm are as follows.

Lemma 4.1. *Consider problem* (23) *with* $\rho_{\text{Basel}} = \rho_{\text{Basel2.5}}$*. Let* $v := -\left(\tilde{R}u^{(j)} + \frac{1}{\sigma_1}\lambda^{(j)}\right)$ *and denote* $v = ((v^{[1]})^\top, (v^{[2]})^\top, \ldots, (v^{[m]})^\top)^\top$*, where* $v^{[s]} \in \mathbb{R}^{n_s}$*,* $s = 1, 2, \ldots, m$*. Let* $(k_{s,1}, k_{s,2}, \ldots, k_{s,n_s})$ *be the permutation of* $(1, 2, \ldots, n_s)$ *such that* $v^{[s]}_{k_{s,1}} \le v^{[s]}_{k_{s,2}} \le \cdots \le v^{[s]}_{k_{s,n_s}}$*,* $s = 1, \ldots, m$*. Let* $p_s := \lceil \alpha n_s \rceil$ *and* $h^{[s]} := (v^{[s]}_{k_{s,1}}, v^{[s]}_{k_{s,2}}, \ldots, v^{[s]}_{k_{s,p_s}})^\top$*,* $s = 1, \ldots, m$*. The optimal solution* x *to* (23) *is given by*

$$x^{[s]}_{k_{s,i}} = \begin{cases} z^{[s]}_i, & \text{if } 1 \le i \le p_s, \\ v^{[s]}_{k_{s,i}}, & \text{otherwise}, \end{cases} \quad i = 1, 2, \ldots, n_s, \tag{29}$$

where $(z^{[1]}, z^{[2]}, \ldots, z^{[m]})$ *is the optimal solution to the following QP problem:*

$$\begin{aligned} \min_{z,\gamma_1,\gamma_2} \quad & \sum_{s=1}^{m} \|z^{[s]} - h^{[s]}\|^2 \\ \text{s.t.} \quad & z_1^{[s]} \le z_2^{[s]} \le \cdots \le z_{p_s}^{[s]}, \quad s = 1, \cdots, m, \\ & \gamma_1 + \gamma_2 \le C_0, \quad z_{p_1}^{[1]} \le \gamma_1, \quad z_{p_{m_1+1}}^{[m_1+1]} \le \gamma_2, \\ & \frac{k}{m_1} \sum_{s=1}^{m_1} z_{p_s}^{[s]} \le \gamma_1, \quad \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} z_{p_s}^{[s]} \le \gamma_2. \end{aligned} \tag{30}$$

Proof. See Appendix A.1. □

Lemma 4.2. *Consider problem* (23) *with* $\rho_{\text{Basel}} = \rho_{\text{Basel3}}$*. Let* v *and* $v^{[s]}$ *be defined as in Lemma 4.1. Let* $x^{[s]}$*,* $s = m_1 + 1, \ldots, m$ *be the optimal solution to the following QP problem:*

$$\begin{aligned} \min_{t,x,z} \quad & \sum_{s=m_1+1}^{m} \|x^{[s]} - v^{[s]}\|^2, \\ \text{s.t.} \quad & t_{m_1+1} + \frac{1}{(1-\alpha)n_{m_1+1}} \sum_{i=1}^{n_{m_1+1}} z_i^{[m_1+1]} \le C_0 \\ & \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} (t_s + \frac{1}{(1-\alpha)n_s} \sum_{i=1}^{n_s} z_i^{[s]}) \le C_0, \\ & z_i^{[s]} \ge 0, z_i^{[s]} \ge x_i^{[s]} - t_s, i = 1, \ldots, n_s, s = m_1 + 1, \ldots, m. \end{aligned} \tag{31}$$

Then the optimal solution to (23) *is given by*

$$x = ((v^{[1]})^\top, (v^{[2]})^\top, \ldots, (v^{[m_1]})^\top, (x^{[m_1+1]})^\top, (x^{[m_1+2]})^\top, \ldots, (x^{[m]})^\top)^\top.$$

Proof. See Appendix A.2. □

Lemma 4.3. *The optimal solution to problem* (24) *with* $\rho = \rho_{\text{Variance}}$ *is*

$$y^{(j+1)} = \sigma_2 \left((\sigma_2 + \frac{2}{n'})I - \frac{2}{(n')^2} \mathbf{1}\mathbf{1}^\top \right)^{-1} w^{(j)} = \left(\sigma_2 + \frac{2}{n'} \right)^{-1} \left(\sigma_2 w^{(j)} + 2 \frac{\mathbf{1}^\top w^{(j)}}{(n')^2} \mathbf{1} \right), \tag{32}$$

where $w^{(j)} = -\left(\tilde{Y} u^{(j)} + \frac{1}{\sigma_2} \pi^{(j)} \right)$.

Proof. See Appendix A.3. □

Lemma 4.4. *Consider problem* (24) *with* $\rho = \rho_{\text{VaR}_\alpha}$*. Define* $w := -(\tilde{Y} u^{(j)} + \frac{1}{\sigma_2} \pi^{(j)}) \in \mathbb{R}^{n'}$ *and* $p' := \lceil \alpha n' \rceil$*. Let* $(k_1, k_2, \ldots, k_{n'})$ *be the permutation of* $(1, 2, \ldots, n')$ *such that* $w_{k_1} \le w_{k_2} \le \cdots \le w_{k_{n'}}$*. Then the optimal solution* y *to* (24) *with* $\rho = \rho_{\text{VaR}_\alpha}$ *is given by*

$$y_{k_i} = \begin{cases} \gamma_{i^*}, & \text{if } i^* \le i \le p', \\ w_{k_i}, & \text{otherwise}, \end{cases} \quad i = 1, 2, \ldots, n', \tag{33}$$

where

$$i^* := \max\{i \mid 1 \le i \le p',\ w_{k_{i-1}} < \gamma_i \le w_{k_i}\},\ w_{k_0} := -\infty,\ \textit{and}\ \gamma_i := \frac{\sigma_2 \sum_{j=i}^{p'} w_{k_j} - 1}{\sigma_2(p' - i + 1)}. \tag{34}$$

Proof. See Appendix A.4. □

Lemma 4.5. *Consider problem* (24) *with* $\rho = \rho_{\mathrm{CVaR}_\alpha}$. *Let* w *be defined as in Lemma* 4.4. *Define*

$$\phi(t, y) := t + \frac{1}{(1-\alpha)n'} \sum_{i=1}^{n'} (y_i - t)_+ + \frac{\sigma_2}{2}\|y - w\|^2, \tag{35}$$

where $x_+ := \max(x, 0)$. *Let* t^* *be the optimal solution to*

$$\min_t\ \phi(t, y(t)),\ \ \textit{s.t.}\ \ \min_{1 \le i \le n'} w_i - c \le t \le \max_{1 \le i \le n'} w_i, \tag{36}$$

where $c = \frac{1}{\sigma_2(1-\alpha)n'}$; $y(t) = (y_1(t), y_2(t), \ldots, y_{n'}(t))^\top$; $y_i(t), i = 1, \ldots, n'$ *are defined by*

$$y_i(t) = \begin{cases} w_i - c, & \textit{if } w_i - c > t, \\ t, & \textit{if } w_i > t \ge w_i - c, \\ w_i, & \textit{otherwise.} \end{cases} \tag{37}$$

Then $y(t^*)$ *is the optimal solution to* (24) *with* $\rho = \rho_{\mathrm{CVaR}_\alpha}$.

Proof. See Appendix A.5. □

4.2 Convergence Analysis

If ρ is ρ_{Variance} or $\rho_{\mathrm{CVaR}_\alpha}$ and ρ_{Basel} is ρ_{Basel3}, problem (16) is convex and the ADM method is ensured to converge to the global solutions theoretically (Hong and Luo, 2012). On the other hand, if ρ is $\rho_{\mathrm{VaR}_\alpha}$ or if ρ_{Basel} is $\rho_{\mathrm{Basel2.5}}$, the convergence of the ADM algorithm to a global optimal solution is not guaranteed due to the non-convexity of $\rho_{\mathrm{VaR}_\alpha}$ or $\rho_{\mathrm{Basel2.5}}$; however, we will show in the following that the limit point of the sequence generated by the ADM algorithm satisfies the first-order optimality conditions of problem (16) under some mild conditions. In addition, numerical experiments suggest that the ADM algorithm seems to converge from any starting point.

We first recall the definition of locally Lipschitz functions.

Definition 4.1. *A function* $f(x) : \mathrm{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ *is Lipschitz near a point* $x_0 \in \mathrm{int}(\mathrm{dom} f)$ *if there exist* $K \ge 0$ *and* $\delta > 0$ *such that* $|f(x) - f(x')| \le K\|x - x'\|$ *for all* $x, x' \in B_\delta(x_0)$, *where* $B_\delta(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \delta\} \subseteq \mathrm{dom} f$. *A function is locally Lipschitz if it is Lipschitz near every point in* $\mathbb{R}^n$. *A function is globally Lipschitz on* $\mathbb{R}^n$ *if there exists a constant* $K \ge 0$ *such that* $|f(x) - f(x')| \le K\|x - x'\|$ *for all* $x, x' \in \mathbb{R}^n$.

We have the following result on the Lipschitz property of risk measures.

Proposition 4.1. *The functions $\rho_{\mathrm{VaR}_\alpha}(x)$, $\rho_{\mathrm{CVaR}_\alpha}(x)$, $\rho_{\mathrm{Basel2.5}}(x)$, and $\rho_{\mathrm{Basel3}}(x)$ defined in Section 3.1 are all globally Lipschitz on $\mathbb{R}^n$.*

Proof. See Appendix B. □

We have the following theorem regarding the optimality of the output of the ADM algorithm.

Theorem 4.1. *Suppose that $\rho(x)$ is locally Lipschitz and $\rho_{\mathrm{Basel}} \in \{\rho_{\mathrm{Basel2.5}}, \rho_{\mathrm{Basel3}}\}$, then the following statements hold.*

(i) (KKT conditions) If u is a local minimizer of (16), then there exits $\eta \geq 0$, such that

$$0 \in -\tilde{Y}^\top \bar{\partial}\rho(y(u)) - \eta \tilde{R}^\top \bar{\partial}\rho_{\mathrm{Basel}}(x(u)) + \mathcal{N}_{\mathcal{U}_{r_0}}(u), \tag{38}$$

$$\eta(\rho_{\mathrm{Basel}}(x(u)) - C_0) = 0, \tag{39}$$

where $\mathcal{N}_{\mathcal{U}_{r_0}}(u)$ is the normal cone to $\mathcal{U}_{r_0}$ at u and $\bar{\partial} f(\cdot)$ denotes the Clarke's generalized gradient of $f(\cdot)$.

(ii) Let $\{(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)})\}$ be a sequence generated by scheme (23)-(27) and assume that $\sum_{j=1}^{\infty} \|\lambda^{(j+1)} - \lambda^{(j)}\|^2 + \|\pi^{(j+1)} - \pi^{(j)}\|^2 < \infty$ and $\{(\lambda^{(j)}, \pi^{(j)})\}$ is bounded. Then, the sequence $\{u^{(j)}\}$ is bounded and any limit point $\bar{u}$ of $\{u^{(j)}\}$ satisfies the first-order optimality conditions (38)-(39).

Proof. See Appendix C. □

By Proposition 4.1, Theorem 4.1 applies to the ADM algorithm with ρ being ρ_{Variance}, $\rho_{\mathrm{VaR}_\alpha}$, or $\rho_{\mathrm{CVaR}_\alpha}$.

4.3 The ADM Algorithm for Solving the Mean-ρ Problems (19) and (20)

The ADM algorithm for solving the mean-ρ problems including the mean-VaR and mean-Basel problems is as follows.

ADM for Solving Problem (19): The augmented Lagrangian function for (19) is defined as

$$\mathcal{L}(x, u, \lambda) := \rho(x) + \lambda^\top (x + \tilde{R}u) + \frac{\sigma}{2}\|x + \tilde{R}u\|^2, \tag{40}$$

where $\sigma > 0$ is the penalty parameter and $\lambda \in \mathbb{R}^n$ is the Lagrangian multiplier. The ADM method is

$$x^{(j+1)} = \arg\min_{x \in \mathbb{R}^n} \ \rho(x) + \frac{\sigma}{2}\|x - v^{(j)}\|^2, \tag{41}$$

$$u^{(j+1)} = \arg\min_{u \in \mathcal{U}_{r_0}} \ \frac{1}{2} u^\top \tilde{R}^\top \tilde{R} u + b^\top u, \tag{42}$$

$$\lambda^{(j+1)} = \lambda^{(j)} + \beta\sigma(x^{(j+1)} + \tilde{R}u^{(j+1)}), \tag{43}$$

where $v^{(j)} = -\left(\tilde{R}u^{(j)} + \frac{1}{\sigma}\lambda^{(j)}\right)$, $b = \tilde{R}^\top(\frac{1}{\sigma}\lambda^{(j)} + x^{(j+1)})$, and $\beta > 0$. For $\rho = \rho_{\text{Variance}}$, ρ_{VaR_α}, and $\rho_{\text{CVaR}_\alpha}$, the subproblem (41) is the same as (24) and hence its solution is given by Lemma 4.3, 4.4, and 4.5, respectively; for $\rho = \rho_{\text{Basel}}$, problem (41) is equivalent to

$$\min_{\tau\in\mathbb{R},x\in\mathbb{R}^n} \tau + \frac{\sigma}{2}\|x - v^{(j)}\|^2, \text{ s.t. } \rho_{\text{Basel}}(x) \leq \tau, \tag{44}$$

whose solution can be obtained in a way similar to that in Lemmas 4.1 and 4.2 except that C_0 in (30) and (31) should be replaced by τ and τ should be added to the objective functions and τ should be included as an additional optimization variable in the minimization. The subproblem (42) can be solved by a standard QP solver.

ADM for Solving Problem (20): The augmented Lagrangian function for (20) is defined as

$$\mathcal{L}_c(x, u, \lambda) := -\mu^\top u + \lambda^\top(x + \tilde{R}u) + \frac{\sigma}{2}\|x + \tilde{R}u\|^2, \tag{45}$$

where $\sigma > 0$ is the penalty parameter and $\lambda \in \mathbb{R}^n$ is the Lagrangian multiplier. The ADM method is

$$x^{(j+1)} = \arg\min_{x\in\mathbb{R}^n} \|x - v^{(j)}\|^2, \text{ s.t. } \rho(x) \leq b_0, \tag{46}$$

$$u^{(j+1)} = \arg\min_{u\in\mathcal{U}} \frac{1}{2}u^\top\tilde{R}^\top\tilde{R}u + b_c^\top u, \tag{47}$$

$$\lambda^{(j+1)} = \lambda^{(j)} + \beta\sigma(x^{(j+1)} + \tilde{R}u^{(j+1)}), \tag{48}$$

where $v^{(j)} = -\left(\tilde{R}u^{(j)} + \frac{1}{\sigma}\lambda^{(j)}\right)$, $b_c = \tilde{R}^\top(\frac{1}{\sigma}\lambda^{(j)} + x^{(j+1)}) - \frac{\mu}{\sigma}$, and $\beta > 0$. For $\rho = \rho_{\text{Variance}}$, (46) is a QP problem; for $\rho = \rho_{\text{CVaR}_\alpha}$, (46) can be formulated as a QP problem by using (12); for $\rho = \rho_{\text{VaR}_\alpha}$, by an argument similar to the proof of Lemma 4.4, the closed-form solution of (46) is given by (33) with γ_{i^*} being replaced by b_0; for $\rho = \rho_{\text{Basel}}$, the solution to (46) can be obtained by Lemmas 4.1 and 4.2. The subproblem (47) is a standard QP problem.

Convergence results similar to Theorem 4.1 can be established for the ADM for models (19) and (20).

5 Numerical Results

In this section, we conduct computational experiments to demonstrate the effectiveness of the ADM method for solving the mean-ρ-Basel model using both simulated and real market data. In particular, we compare the performance of ADM method with that of MIP/QP/LP solvers in CPLEX 12.4. The numerical results suggest that the ADM method is promising in generating solutions of high quality to the model in reasonable computational time.

5.1 Data Description

In our experiments, the real market data and simulated data sets are generated as follows.

- **S&P 500 Data Set.** The S&P 500 data set comprises the daily returns of 359 stocks that have ever been included in the S&P 500 index and do not have missing data during the following specified time periods. Let $t_0 = 03/01/2012$. For $s = 1, \ldots, 60$, $\tilde{R}_{SP}^{[s]}$ denotes the trailing five-year daily returns of the stocks on day $t_0 - s + 1$ (i.e., the daily returns of the stocks during the period from day $t_0 - s - 2058$ to day $t_0 - s + 1$). Let $l = 06/01/2007$ and $u = 06/01/2009$. For $s = 61, \ldots, 120$, $\tilde{R}_{SP}^{[s]}$ is defined as the daily returns of the stocks during the stressed period from day $l + 120 - s$ to day $u - s + 61$. Then the S&P data matrix $\tilde{R}_{SP}$ is defined from $\tilde{R}_{SP}^{[s]}, s = 1, \ldots, 120$ by Eq. (7).

- **Simulated Data.** We simulate the prices of 350 stocks based on a multi-dimensional version of the double-exponential jump diffusion model (Kou, 2002):

$$\frac{dS_i(t)}{S_i(t-)} = \mu_i dt + \sigma_i dW_i(t) + d\left(\sum_{k=1}^{N_i(t)} (e^{V_{ik}} - 1)\right), i = 1, \ldots, n, \tag{49}$$

  where $W_1(t), \ldots, W_n(t)$ are n correlated Brownian motions with $dW_i(t)dW_j(t) = \rho_{ij}dt$; $N_i(t)$ is a Poisson process with intensity λ_i; $N_i(t)$ is independent of $N_j(t)$ for $i \neq j$; $\{V_{i1}, V_{i2}, \ldots\}$ are i.i.d. log jump sizes with a double-exponential probability density function $f_i(x) = p_i \eta_{iu} e^{-\eta_{iu}x} 1_{\{x \geq 0\}} + (1 - p_i)\eta_{id} e^{\eta_{id}x} 1_{\{x<0\}}$; V_{ik} and V_{jl} are independent for $i \neq j$; and the Brownian motions, Poisson processes, and jump sizes are mutually independent. The stock returns generated in the above model have the same tail heaviness as those generated by the negative exponential tail model considered in Lim, Shanthikumar, and Vahn (2011). Two sets of parameters $\{\mu_i, \sigma_i, \lambda_i, p_i, \eta_{iu}, \eta_{id}, \rho_{ij}\}$ are used to simulate stock returns under normal and stressed market conditions, respectively; these parameters are estimated from the historical data of some large-cap stocks during normal and stressed market conditions, respectively. Δt is set to be $1/252$ (one day).

5.2 Parameter Settings of the ADM and MIP

Our method is implemented in MATLAB. All the experiments were performed on a Dell Precision Workstation T5500 with Intel Xeon CPU E5620 at 2.40GHz and 12GB of memory running Ubuntu 12.04 and MATLAB 2011b. All the quadratic programming subproblems in the ADM method are solved using the QP solvers in CPLEX 12.4 with Matlab interface; and the mixed integer programming (MIP) reformulations of the asset allocation models are solved using the MIP solvers in CPLEX 12.4. In our test, the parameters σ and β in (23)-(26) are set to be 10^{-3} and 0.1, respectively. The initial Lagrangian multipliers are $\lambda^{(0)} = 0$ and $\pi^{(0)} = 0$. The method is terminated if either

Table 1: The number of binary variables, continuous variables, and linear constraints in the MIP/QP formulation of the mean-variance-Basel problems.

	$\rho_{\text{Basel2.5}}(x(u)) \le C_0$			$\rho_{\text{Basel3}}(x(u)) \le C_0$		
d	binary	continuous	constraints	binary	continuous	constraints
100	58020	4601	62526	–	32319	32163
150	58020	4651	62526	–	32369	32163
200	58020	4701	62526	–	32419	32163
250	58020	4751	62526	–	32469	32163
300	58020	4801	62526	–	32519	32163
350	58020	4851	62526	–	32569	32163

$\|x^{(j+1)} + \tilde{R}u^{(j+1)}\|^2 + \|y^{(j+1)} + \tilde{Y}u^{(j+1)}\|^2 \le 10^{-8}$, $\frac{\|u^{(j+1)}-u^{(j)}\|}{\max(1,\|u^{(j)}\|)} \le 10^{-4}$, or the number of iterations has reached an upper bound of 2000. The default setting of the MIP solver in CPLEX 12.4 is used. The maximum CPU time limit for all solvers is set to 3600 seconds.

5.3 Comparing ADM with MIP/QP on the Mean-Variance-Basel Model

In this subsection, we evaluate the performance of the ADM on the mean-variance-Basel problems:

$$\begin{aligned} \min_{u\in\mathcal{U}_{r_0}} \ & \rho_{\text{Variance}}(y(u)) \\ \text{s.t.} \ & \rho_{\text{Basel2.5}}(x(u)) \le C_0, \end{aligned} \quad \text{and} \quad \begin{aligned} \min_{u\in\mathcal{U}_{r_0}} \ & \rho_{\text{Variance}}(y(u)) \\ \text{s.t.} \ & \rho_{\text{Basel3}}(x(u)) \le C_0. \end{aligned} \tag{50}$$

The mean-variance-Basel2.5 problem can be solved using the MIP method, and the mean-variance-Basel3 problem is a QP problem that can be solved using the QP solver in CPLEX 12.4.

We compare the ADM with the MIP/QP methods for the two problems, respectively, for different numbers of stocks $d \in \{100, 150, 200, 250, 300, 350\}$ using real market and simulated data. For the real market data, $\tilde{R}$ is defined to be a submatrix of $\tilde{R}_{SP}$ consisting of d columns of $\tilde{R}_{SP}$ that are randomly selected. The mean μ used for defining $\mathcal{U}_{r_0}$ is set as the sample mean of $\tilde{R}$. The prescribed return level r_0 is set to be the 80% quantile of the cross-sectional expected returns of the d stocks. $\tilde{Y}$ is obtained by deleting the duplicated rows in $\tilde{R}$. The parameters in (13) are set at $\alpha = 0.99$, $k = 3$, and $\ell = 3$; those in (15) are set at $\alpha = 0.98$ and $\ell = 6$. C_0 is set at 0.2. Table 1 reports the numbers of binary variables, continuous variables, and linear constraints, denoted by "binary," "continuous," and "constraints," respectively, of the two problems in (50).

The optimal objective value $\rho_{\text{Variance}}(y(u))$ obtained and the CPU time used by the ADM and MIP/QP methods for the simulated and real market data are presented in Figures 1 and 2, respectively. These values, as well as $\rho_{\text{Basel2.5}}(x(u))$ and $\rho_{\text{Basel3}}(x(u))$, are reported

Table 2: The numerical results of solving the mean-variance-Basel problems with simulated data using the ADM and the MIP/QP methods.

stocks	$\text{ADM}_{\text{Basel2.5}\leq\text{C}_0}$			$\text{MIP}_{\text{Basel2.5}\leq\text{C}_0}$			$\text{ADM}_{\text{Basel3}\leq\text{C}_0}$			$\text{QP}_{\text{Basel3}\leq\text{C}_0}$		
d	ρ_{Variance}	time	$\rho_{\text{Basel2.5}}$	ρ_{Variance}	time	$\rho_{\text{Basel2.5}}$	ρ_{Variance}	time	ρ_{Basel3}	ρ_{Variance}	time	ρ_{Basel3}
100	0.4020	108	0.146	0.6399	3600	0.158	0.4088	266	0.200	0.8277	3603	0.144
150	0.4538	120	0.164	0.5029	3631	0.171	0.4701	268	0.200	0.4702	572	0.200
200	0.4054	117	0.164	0.4664	3605	0.147	0.4087	267	0.200	0.4087	567	0.200
250	0.4090	127	0.161	0.4895	3630	0.153	0.4226	268	0.200	0.4227	653	0.200
300	0.3437	128	0.151	0.4356	3613	0.154	0.3561	288	0.200	0.3562	671	0.200
350	0.3428	132	0.156	0.4169	3605	0.163	0.3543	296	0.200	0.3544	717	0.200

in Tables 2 and 3. We can observe that the ADM obtains a better objective value of ρ_{Variance} and is faster than the MIP/QP methods.

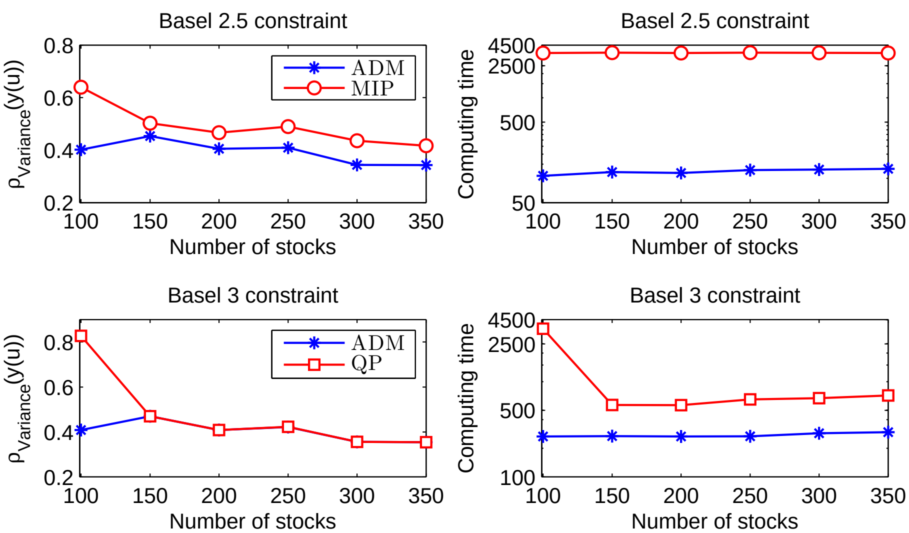

Figure 1: Comparing the ADM with the MIP for the mean-variance-Basel2.5 problem and comparing the ADM with the QP for the mean-variance-Basel3 problem for different numbers of stocks using simulated data. CPU time is expressed in seconds.

5.4 Comparing ADM with MIP/LP on the Mean-CVaR-Basel Model

In this subsection, we evaluate the performance of the ADM on the mean-CVaR-Basel problems:

$$\begin{aligned}\min_{u\in\mathcal{U}_{r_0}} \quad & \rho_{\text{CVaR}_\alpha}(y(u)) \\ \text{s.t.} \quad & \rho_{\text{Basel2.5}}(x(u)) \leq C_0,\end{aligned} \quad \text{and} \quad \begin{aligned}\min_{u\in\mathcal{U}_{r_0}} \quad & \rho_{\text{CVaR}_\alpha}(y(u)) \\ \text{s.t.} \quad & \rho_{\text{Basel3}}(x(u)) \leq C_0.\end{aligned} \tag{51}$$

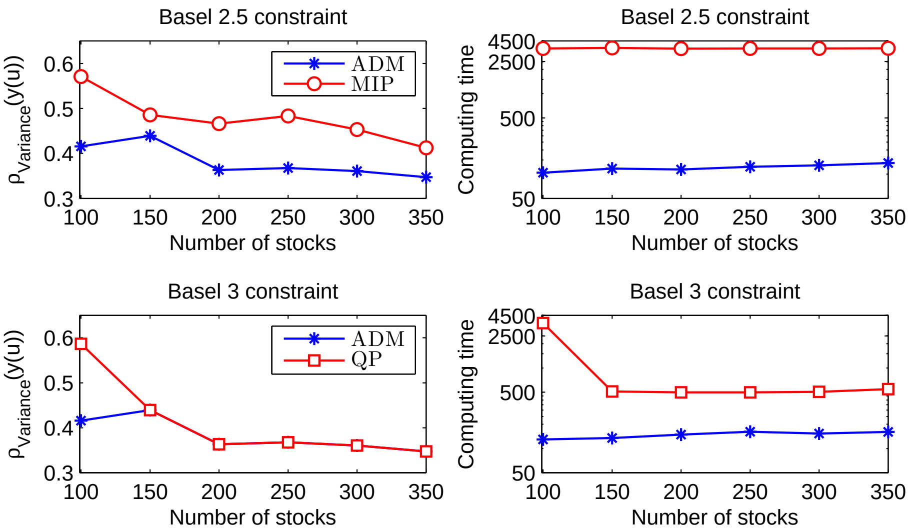

Figure 2: Comparing the ADM with the MIP for the mean-variance-Basel2.5 problem and comparing the ADM with the QP for the mean-variance-Basel3 problem for different numbers of stocks using real market data. CPU time is expressed in seconds.

The mean-CVaR-Basel2.5 problem can be solved using the MIP method, and the mean-CVaR-Basel3 problem can be formulated as a LP problem and solved using the dual simplex (LP) solver in CPLEX 12.4.

We compare the ADM with the MIP/LP methods for the two problems, respectively. The setup of the experiments is the same as in Section 5.3. Hence, the numbers of binary variables and linear constraints are the same as those in Table 1, and the number of continuous variables is equal to that in Table 1 plus one.

The optimal objective value $\rho_{\text{CVaR}_\alpha}(y(u))$ obtained and the CPU time used by the ADM and MIP/LP methods for the simulated and real market data are presented in Figures 3 and 4, respectively. These values, as well as $\rho_{\text{Basel2.5}}(x(u))$ and $\rho_{\text{Basel3}}(x(u))$, are reported in

Table 3: The numerical results of solving the mean-variance-Basel problems with real market data using the ADM and the MIP/QP methods.

stocks	$\text{ADM}_{\text{Basel2.5}\le\text{C}_0}$			$\text{MIP}_{\text{Basel2.5}\le\text{C}_0}$			$\text{ADM}_{\text{Basel3}\le\text{C}_0}$			$\text{QP}_{\text{Basel3}\le\text{C}_0}$		
d	ρ_{Variance}	time	$\rho_{\text{Basel2.5}}$	ρ_{Variance}	time	$\rho_{\text{Basel2.5}}$	ρ_{Variance}	time	ρ_{Basel3}	ρ_{Variance}	time	ρ_{Basel3}
100	0.4157	104	0.145	0.5710	3627	0.148	0.4157	130	0.157	0.5867	3610	0.122
150	0.4393	118	0.155	0.4859	3698	0.136	0.4393	135	0.145	0.4393	510	0.145
200	0.3633	114	0.133	0.4662	3619	0.135	0.3633	149	0.141	0.3633	497	0.141
250	0.3678	124	0.145	0.4832	3630	0.139	0.3678	162	0.152	0.3678	497	0.152
300	0.3608	129	0.138	0.4530	3638	0.148	0.3608	153	0.144	0.3608	505	0.144
350	0.3473	137	0.137	0.4125	3649	0.138	0.3473	161	0.147	0.3473	547	0.147

Table 4: The numerical results of solving the mean-CVaR-Basel problems with simulated data using the ADM and the MIP/LP methods.

stocks	$\text{ADM}_{\text{Basel2.5}\le C_0}$			$\text{MIP}_{\text{Basel2.5}\le C_0}$			$\text{ADM}_{\text{Basel3}\le C_0}$			$\text{LP}_{\text{Basel3}\le C_0}$		
d	$\rho_{\text{CVaR}_\alpha}$	time	$\rho_{\text{Basel2.5}}$	$\rho_{\text{CVaR}_\alpha}$	time	$\rho_{\text{Basel2.5}}$	$\rho_{\text{CVaR}_\alpha}$	time	ρ_{Basel3}	$\rho_{\text{CVaR}_\alpha}$	time	ρ_{Basel3}
100	0.0259	108	0.135	0.0258	583	0.134	0.0259	252	0.194	0.0258	6	0.197
150	0.0273	115	0.151	0.0272	866	0.151	0.0273	268	0.198	0.0272	7	0.194
200	0.0259	116	0.141	0.0258	1099	0.143	0.0259	252	0.170	0.0258	7	0.172
250	0.0262	118	0.146	0.0261	1365	0.142	0.0262	264	0.184	0.0261	9	0.185
300	0.0243	124	0.132	0.0243	1640	0.132	0.0243	276	0.180	0.0243	28	0.184
350	0.0243	133	0.134	0.0242	1981	0.131	0.0243	284	0.178	0.0242	19	0.181

Tables 4 and 5. We can observe that (i) the ADM method is faster than the MIP method for the mean-CVaR-Basel2.5 problem but is slower than the LP method for the mean-CVaR-Basel3 problem; (ii) the optimal objective value $\rho_{\text{CVaR}_\alpha}(y(u))$ obtained by the ADM and the MIP/LP are almost the same. In fact, the largest absolute value of the relative difference ("rel.dif") of $\rho_{\text{CVaR}_\alpha}$ between that obtained using the ADM and that obtained using the MIP/LP is 0.39%, where "rel.dif" is defined by rel.dif $:= (\rho_{\text{CVaR}_\alpha}(y(u_{\text{ADM}})) - \rho_{\text{CVaR}_\alpha}(y(u_{\text{MIP/LP}})))/\rho_{\text{CVaR}_\alpha}(y(u_{\text{MIP/LP}}))$.

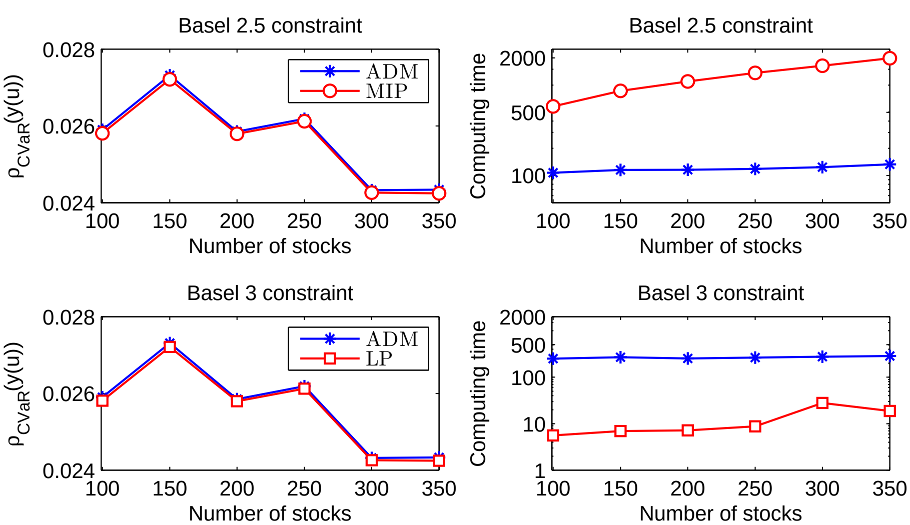

Figure 3: Comparing the ADM with the MIP for the mean-CVaR-Basel2.5 problem and comparing the ADM with the LP for the mean-CVaR-Basel3 problem for different numbers of stocks using simulated data. CPU time is expressed in seconds.

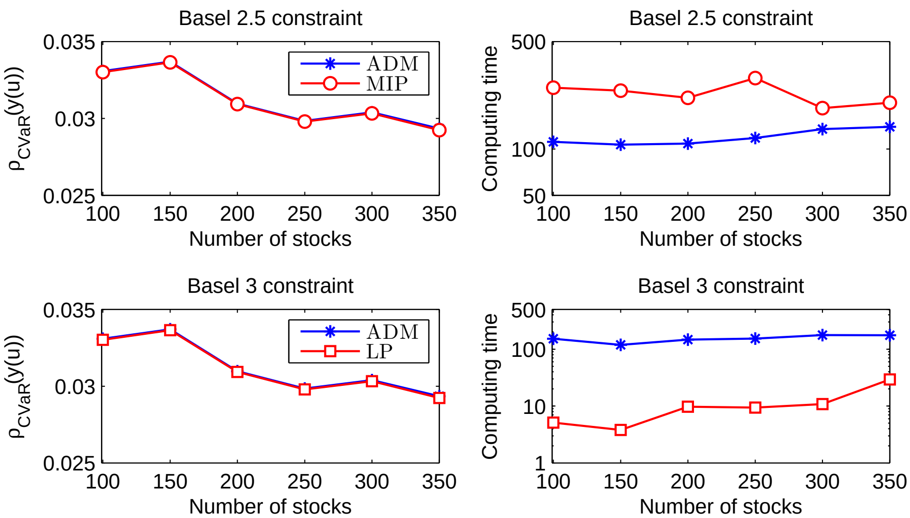

Figure 4: Comparing the ADM with the MIP for the mean-CVaR-Basel2.5 problem and comparing the ADM with the LP for the mean-CVaR-Basel3 problem for different numbers of stocks using real market data. CPU time is expressed in seconds.

5.5 Comparing ADM with MIP on the Mean-VaR-Basel Model

In this subsection, we compare the performance of the ADM with that of the MIP on the mean-VaR-Basel models:

$$\begin{aligned} \min_{u\in\mathcal{U}_{r_0}} \ & \rho_{\text{VaR}_\alpha}(y(u)) \\ \text{s.t.} \ & \rho_{\text{Basel2.5}}(x(u)) \le C_0, \end{aligned} \quad \text{and} \quad \begin{aligned} \min_{u\in\mathcal{U}_{r_0}} \ & \rho_{\text{VaR}_\alpha}(y(u)) \\ \text{s.t.} \ & \rho_{\text{Basel3}}(x(u)) \le C_0. \end{aligned} \tag{52}$$

The setup of the experiments is the same as that in Section 5.3. Table 6 reports the number of binary variables, continuous variables, and linear constraints in the MIP formulation of the problems in (52).

Table 5: The numerical results of solving the mean-CVaR-Basel problems with real market data using the ADM and the MIP/LP methods.

stocks	$\text{ADM}_{\text{Basel2.5}\le \text{C}_0}$			$\text{MIP}_{\text{Basel2.5}\le \text{C}_0}$			$\text{ADM}_{\text{Basel3}\le \text{C}_0}$			$\text{LP}_{\text{Basel3}\le \text{C}_0}$		
d	$\rho_{\text{CVaR}_\alpha}$	time	$\rho_{\text{Basel2.5}}$	$\rho_{\text{CVaR}_\alpha}$	time	$\rho_{\text{Basel2.5}}$	$\rho_{\text{CVaR}_\alpha}$	time	ρ_{Basel3}	$\rho_{\text{CVaR}_\alpha}$	time	ρ_{Basel3}
100	0.0331	111	0.141	0.0330	251	0.143	0.0331	153	0.156	0.0330	5	0.160
150	0.0337	107	0.133	0.0336	240	0.134	0.0337	119	0.143	0.0336	4	0.139
200	0.0310	109	0.132	0.0309	215	0.139	0.0310	147	0.141	0.0309	10	0.147
250	0.0298	118	0.142	0.0298	289	0.145	0.0298	153	0.151	0.0298	9	0.154
300	0.0304	135	0.140	0.0303	184	0.141	0.0304	177	0.147	0.0303	11	0.153
350	0.0293	140	0.140	0.0292	200	0.142	0.0293	175	0.151	0.0292	29	0.150

Table 6: The number of binary variables, continuous variables, and linear constraints in the MIP formulation of the mean-VaR-Basel problems.

	$\rho_{\text{Basel2.5}}(x(u)) \le C_0$			$\rho_{\text{Basel3}}(x(u)) \le C_0$		
d	binary	continuous	constraints	binary	continuous	constraints
100	62399	223	62527	4379	27941	32164
150	62399	273	62527	4379	27991	32164
200	62399	323	62527	4379	28041	32164
250	62399	373	62527	4379	28091	32164
300	62399	423	62527	4379	28141	32164
350	62399	473	62527	4379	28191	32164

Table 7: The numerical results obtained when solving the mean-VaR-Basel problems with simulated data using the ADM and the MIP methods.

stocks	$\text{ADM}_{\text{Basel2.5}\le \text{C}_0}$			$\text{MIP}_{\text{Basel2.5}\le \text{C}_0}$			$\text{ADM}_{\text{Basel3}\le \text{C}_0}$			$\text{MIP}_{\text{Basel3}\le \text{C}_0}$		
d	ρ_{VaR}	time	$\rho_{\text{Basel2.5}}$	ρ_{VaR}	time	$\rho_{\text{Basel2.5}}$	ρ_{VaR}	time	ρ_{Basel3}	ρ_{VaR}	time	ρ_{Basel3}
100	0.0214	174	0.151	0.0254	3602	0.158	0.0213	361	0.200	0.0207	3601	0.200
150	0.0231	162	0.174	0.0265	3602	0.171	0.0233	397	0.200	0.0225	3601	0.200
200	0.0219	160	0.170	0.0202	3607	0.161	0.0214	390	0.200	0.0203	3601	0.200
250	0.0210	175	0.166	0.0239	3604	0.151	0.0217	402	0.200	0.0206	3602	0.200
300	0.0195	180	0.162	0.0243	3611	0.153	0.0201	422	0.200	0.0192	3605	0.200
350	0.0197	192	0.163	0.0236	3612	0.144	0.0200	428	0.200	0.0194	3609	0.200

The optimal objective value $\rho_{\text{VaR}_\alpha}(y(u))$ obtained and the CPU time used by the ADM and the MIP methods for the simulated and real market data are presented in Figures 5 and 6, respectively. These values, as well as $\rho_{\text{Basel2.5}}(x(u))$ and $\rho_{\text{Basel3}}(x(u))$, are reported in Tables 7 and 8. The figures and tables show that the ADM is a very good alternative to the MIP for the mean-VaR-Basel problems because: (i) The ADM is much faster than the MIP. (ii) For the mean-VaR-Basel2.5 problem, the optimal objective value ρ_{VaR_α} computed by the ADM is smaller than that computed by the MIP except in two cases; in fact, the relative difference of ρ_{VaR} between the ADM and the MIP, which is defined by $(\rho_{\text{VaR}}(y(u_{\text{ADM}})) - \rho_{\text{VaR}}(y(u_{\text{MIP}})))/\rho_{\text{VaR}}(y(u_{\text{MIP}}))$, is in the range of $[-19.65\%, 8.07\%]$, which shows that the ADM may be slightly inferior to the MIP in some cases but can be significantly preferable to the MIP in other cases. (iii) For the mean-VaR-Basel3 problem, the relative difference of ρ_{VaR} between the ADM and the MIP is in the range of $[-23.18\%, 5.44\%]$, which shows that overall the ADM achieves better objective value than the MIP.

6 Conclusions

A major change in financial regulations after the recent financial crisis is that financial institutions are now required to meet more stringent regulatory capital requirements than

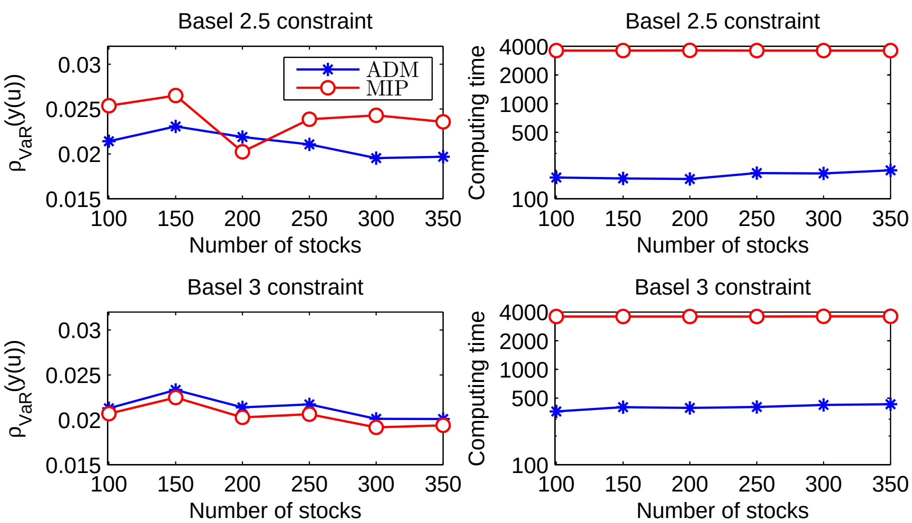

Figure 5: Comparing the ADM with the MIP for the mean-VaR-Basel problems for different numbers of stocks using simulated data. CPU time is expressed in seconds.

Table 8: The numerical results obtained when solving the mean-VaR-Basel problems with real market data using the ADM and the MIP methods.

stocks	$\mathrm{ADM}_{\mathrm{Basel2.5} \le \mathrm{C}_0}$			$\mathrm{MIP}_{\mathrm{Basel2.5} \le \mathrm{C}_0}$			$\mathrm{ADM}_{\mathrm{Basel3} \le \mathrm{C}_0}$			$\mathrm{MIP}_{\mathrm{Basel3} \le \mathrm{C}_0}$		
d	ρ_{VaR}	time	$\rho_{\mathrm{Basel2.5}}$	ρ_{VaR}	time	$\rho_{\mathrm{Basel2.5}}$	ρ_{VaR}	time	ρ_{Basel3}	ρ_{VaR}	time	ρ_{Basel3}
100	0.0246	148	0.138	0.0238	3602	0.143	0.0245	247	0.154	0.0237	3601	0.178
150	0.0249	167	0.146	0.0284	3602	0.160	0.0248	244	0.153	0.0259	3601	0.195
200	0.0228	174	0.129	0.0268	3602	0.159	0.0228	269	0.142	0.0222	3602	0.154
250	0.0233	194	0.134	0.0256	3606	0.153	0.0233	300	0.164	0.0248	3602	0.171
300	0.0224	193	0.133	0.0265	3606	0.158	0.0224	283	0.152	0.0291	3602	0.193
350	0.0227	208	0.131	0.0270	3609	0.146	0.0228	302	0.149	0.0293	3610	0.197

previously. It has been estimated that the capital requirement for a large bank's trading book under the Basel 2.5 Accord on average *more than doubles* that under the Basel II Accord. The significantly higher capital requirement makes it more important for banks to take into account the capital constraint when they construct their investment portfolios. In this paper, we propose a new asset allocation model, called the "mean-ρ-Basel" model, that incorporates the Basel Accord capital requirements as one of the constraints. In this model, the capital requirement is measured using the Basel 2.5 and Basel III risk measures imposed by regulators; the risk level of the portfolio is measured by ρ, such as variance, VaR, and CVaR that can be freely chosen by the portfolio manager.

The complexity of the Basel 2.5 and Basel III risk measures, which involve risk measurement under multiple scenarios, including stressed scenarios, poses significant computational challenges to the proposed asset allocation problem due to its inherent non-convexity

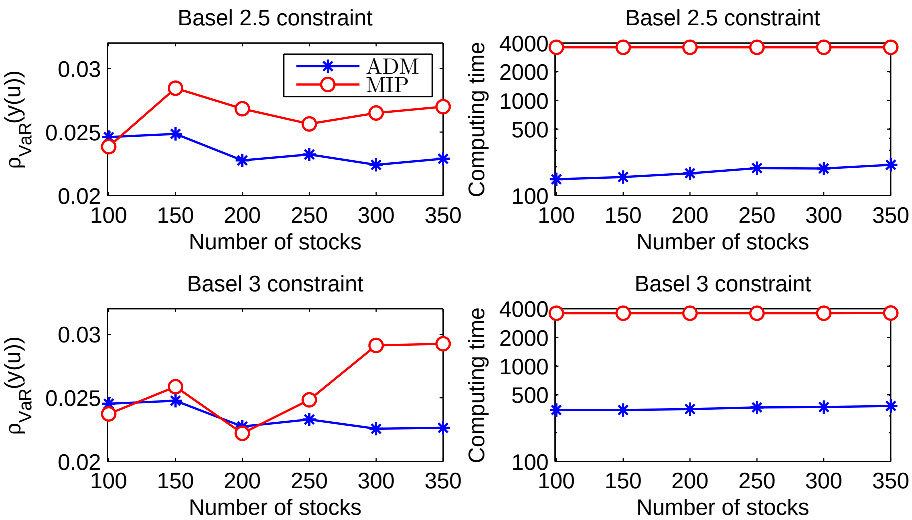

Figure 6: Comparing the ADM with the MIP for the mean-VaR-Basel problems for different numbers of stocks using real market data. CPU time is expressed in seconds.

and non-smoothness. We propose an unified algorithm based on the alternating direction augmented Lagrangian method to solve the mean-ρ-Basel model and classical mean-ρ model. The method is very simple and easy to implement; it reduces the original problem to one-dimensional optimization or convex quadratic programming subproblems that may even have closed-form solutions; hence, it is capable of solving large-scale problems that are difficult to solve using many other methods. For non-convex cases of the mean-ρ-Basel model, we establish the first-order optimality of the limit points of the sequence generated by the method under some mild conditions. Extensive numerical results suggest that our method is promising for finding high-quality approximate optimal solutions, especially in non-convex cases.

A Proof of Lemmas in Section 4.1

A.1 Proof of Lemma 4.1

Proof. Since $\rho_{\text{Basel}} = \rho_{\text{Basel2.5}}$, problem (23) is equivalent to

$$\begin{aligned} \min_x \ & \phi(x) := \sum_{s=1}^m \|x^{[s]} - v^{[s]}\|^2 \\ \text{s.t.} \ & \max\left\{x^{[1]}_{(p_1)}, \frac{k}{m_1}\sum_{s=1}^{m_1} x^{[s]}_{(p_s)}\right\} + \max\left\{x^{[m_1+1]}_{(p_{m_1+1})}, \frac{\ell}{m_2}\sum_{s=m_1+1}^{m} x^{[s]}_{(p_s)}\right\} \le C_0. \end{aligned} \tag{53}$$

Without loss of generality, assume that $v_1^{[s]} \le v_2^{[s]} \le \cdots \le v_{n_s}^{[s]}$, $s = 1, \dots, m$. Then, $(k_{s,1}, k_{s,2}, \dots, k_{s,n_s}) = (1, 2, \dots, n_s)$. Let x be an optimal solution to (53). If $x_i^{[s]} > x_j^{[s]}$ for some $i < j$, then since $v_i^{[s]} \le v_j^{[s]}$, switching the values of $x_i^{[s]}$ and $x_j^{[s]}$ will maintain the feasibility of x without increasing $\phi(x)$. Thus, we can obtain an optimal solution x that satisfies $x_1^{[s]} \le x_2^{[s]} \le \cdots \le x_{n_s}^{[s]}$. In addition, it must hold that $x_i^{[s]} \le v_i^{[s]}$ for all i; otherwise, if $x_i^{[s]} > v_i^{[s]}$ for some i, then setting $x_i^{[s]} = v_i^{[s]}$ will maintain the feasibility of x but strictly reduce $\phi(x)$. Furthermore, it must hold that $x_j^{[s]} = v_j^{[s]}$ for all $j > p_s$; otherwise, if there is some $j > p_s$ such that $x_j^{[s]} < v_j^{[s]}$, setting $x_j^{[s]} = v_j^{[s]}$ will maintain the feasibility of x but strictly reduce $\phi(x)$. Therefore, problem (53) is equivalent to

$$
\begin{aligned}
\min_x \ & \phi(x) = \sum_{s=1}^{m} \|x^{[s]} - v^{[s]}\|^2 \\
\text{s.t.} \ & x_1^{[s]} \le x_2^{[s]} \le \cdots \le x_{p_s}^{[s]}, \ s = 1, \dots, m, \\
& x_j^{[s]} = v_j^{[s]}, \ j = p_s + 1, \dots, n_s, \ s = 1, \dots, m, \\
& \max\left\{ x_{p_1}^{[1]}, \frac{k}{m_1} \sum_{s=1}^{m_1} x_{p_s}^{[s]} \right\} + \max\left\{ x_{p_{m_1+1}}^{[m_1+1]}, \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} x_{p_s}^{[s]} \right\} \le C_0.
\end{aligned}
\tag{54}
$$

Hence, the optimal solution x is given by (29). □

A.2 Proof of Lemma 4.2

Proof. The problem (23) with $\rho_{\text{Basel}} = \rho_{\text{Basel3}}$ is equivalent to

$$
\begin{aligned}
\min_x \ & \sum_{s=1}^{m} \|x^{[s]} - v^{[s]}\|^2 \\
\text{s.t.} \ & \max\left\{ \rho_{\text{CVaR}_\alpha}(x^{[m_1+1]}), \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \rho_{\text{CVaR}_\alpha}(x^{[s]}) \right\} \le C_0.
\end{aligned}
\tag{55}
$$

Let $x = ((x^{[1]})^\top, (x^{[2]})^\top, \dots, (x^{[m]})^\top)^\top$ be the optimal solution to (55). Then apparently $x^{[s]} = v^{[s]}$, for $s = 1, 2, \dots, m_1$. By (12),

$$
\rho_{\text{CVaR}_\alpha}(x^{[s]}) = \min_{t \in \mathbb{R}} \ t + \frac{1}{(1-\alpha) n_s} \sum_{i=1}^{n_s} (x_i^{[s]} - t)_+. \tag{56}
$$

Then using (56), it is easy to show that $((x^{[m_1+1]})^\top, (x^{[m_1+2]})^\top, \dots, (x^{[m]})^\top)^\top$ is an optimal solution to (31), which completes the proof. □

A.3 Proof of Lemma 4.3.

Proof. The subproblem (24) is equivalent to

$$y^{(j+1)} = \arg\min_{y} \ \rho(y) + \frac{\sigma_2}{2}\|y - w^{(j)}\|^2, \text{ where } w^{(j)} = -\left(\tilde{Y}u^{(j)} + \frac{1}{\sigma_2}\pi^{(j)}\right). \tag{57}$$

The result follows by using the definition of ρ_{Variance} given in (9) and Sherman–Morrison–Woodbury formula. □

A.4 Proof of Lemma 4.4.

Proof. The problem (24) with $\rho = \rho_{\text{VaR}_\alpha}$ becomes

$$\min_{y} \ \psi(y) = y_{(p')} + \frac{\sigma_2}{2}\|y - w\|^2, \text{ where } p' := \lceil \alpha n' \rceil. \tag{58}$$

Without loss of generality, assume that $w_1 \leq w_2 \leq \ldots \leq w_{n'}$. Then $(k_1, k_2, \ldots, k_{n'}) = (1, 2, \ldots, n')$. Let y be an optimal solution of (58). If $y_i > y_j$ for some $i < j$, then since $w_i \leq w_j$, switching the values of y_i and y_j will not increase $\psi(y)$. Thus, we can obtain an optimal solution y that satisfies $y_1 \leq y_2 \leq \cdots \leq y_{n'}$. In addition, the optimal solution y must satisfy that $y_i \leq w_i$ for all i; otherwise, if $y_i > w_i$ for some i, then setting $y_i = w_i$ will strictly reduce $\psi(y)$. Furthermore, it must hold that $y_j = w_j$ for all $j = p'+1, p'+2, \ldots, n'$; otherwise, if there is some $j > p'$ such that $y_j < w_j$, setting $y_j = w_j$ will strictly reduce $\psi(y)$. Therefore, problem (58) is equivalent to

$$\begin{aligned} \min_{y} \ & y_{p'} + \frac{\sigma_2}{2}\sum_{i=1}^{p'}(y_i - w_i)^2 \\ \text{s.t.} \ & y_i \leq y_{p'}, \ i = 1, \cdots, p'-1, \\ & y_j = w_j, \ j = p'+1, p'+2, \ldots, n'. \end{aligned} \tag{59}$$

The KKT conditions of (59) are

$$y_i \leq y_{p'}, \ i = 1, \ldots, p'-1, \tag{60}$$

$$\sigma_2(y_i - w_i) + \bar{\pi}_i = 0, \ i = 1, \ldots, p'-1, \tag{61}$$

$$\bar{\pi}_i(y_{p'} - y_i) = 0, \ i = 1, \ldots, p'-1, \tag{62}$$

$$\sigma_2(y_{p'} - w_{p'}) + 1 - \sum_{j=1}^{p'-1}\bar{\pi}_j = 0, \tag{63}$$

$$\bar{\pi}_i \geq 0, \ i = 1, \ldots, p'-1. \tag{64}$$

Since problem (59) is convex, the KKT conditions are also sufficient for the optimality of y. The equations (61) and (62) imply that for each $i = 1, \ldots, p'-1$, either $y_i = w_i$ (if $\bar{\pi}_i = 0$) or $y_i = y_{p'}$ (if $\bar{\pi}_i > 0$). Since $y_1 \leq \cdots \leq y_{p'}$, it follows that there exists

$1 \le i^* \le p'$ such that $y_j = w_j$ for $j < i^*$, $y_j = y_{p'}$ for $j \ge i^*$, and $w_{i^*-1} < y_{p'}$. Then by (63), we have $y_{p'} = \gamma_{i^*}$, where γ_i is defined in (34). It follows from (61) and (64) that $y_j \le w_j$, $j = 1, \ldots, p'$. Hence, we have $\gamma_{i^*} = y_{i^*} \le w_{i^*}$. Therefore, i^* should satisfy $w_{i^*-1} < \gamma_{i^*} \le w_{i^*}$, which completes the proof. ☐

A.5 Proof of Lemma 4.5

Proof. With $\rho = \rho_{\mathrm{CVaR}_\alpha}$, it follows from (56) that problem (24) is equivalent to

$$\min_{t,y} \ \phi(t,y) = t + \frac{1}{(1-\alpha)n'} \sum_{i=1}^{n'} (y_i - t)_+ + \frac{\sigma_2}{2} \|y - w^{(j)}\|^2, \tag{65}$$

where $x_+ := \max(x, 0)$. For any fixed t, the optimal y that minimizes $\phi(t,y)$ is $y(t)$ defined in (37). Hence, the result follows. ☐

B Proof of Proposition 4.1.

Proof. We first show that for any fixed $1 \le p \le n$, $f(x) := x_{(p)}$ is globally Lipschitz. For any given $x \in \mathbb{R}^n$, define

$$L_{x_{(p)}} := \{i \mid x_i < x_{(p)}\}, \ E_{x_{(p)}} := \{i \mid x_i = x_{(p)}\}, \ G_{x_{(p)}} := \{i \mid x_i > x_{(p)}\}. \tag{66}$$

For any given $y, z \in \mathbb{R}^n$, without loss of generality, assume that $y_{(p)} \le z_{(p)}$. It follows from the definition of $L_{y_{(p)}}$ and $E_{y_{(p)}}$ that the number of elements of $L_{y_{(p)}} \cup E_{y_{(p)}}$ is strictly larger than that of $L_{z_{(p)}}$. Therefore, the set $I := (L_{y_{(p)}} \cup E_{y_{(p)}}) \cap (E_{z_{(p)}} \cup G_{z_{(p)}})$ is not empty. Choose any $i \in I$. Then $y_i \le y_{(p)}$ and $z_i \ge z_{(p)}$. Hence, $|y_{(p)} - z_{(p)}| = z_{(p)} - y_{(p)} \le z_i - y_i \le \|y - z\|$, which establishes that $\rho_{\mathrm{VaR}_\alpha}(x)$ is globally Lipschitz.

Using the inequality $|\max(a,b) - \max(c,d)| \le |a-c| + |b-d|$ for $\forall a, b, c, d \in \mathbb{R}$, it can be shown that the maximum of two globally Lipschitz functions is also globally Lipschitz. Since $\rho_{\mathrm{CVaR}_\alpha}$, $\rho_{\mathrm{Basel2.5}}$, and ρ_{Basel3} are all finite linear combination of (maximum of) globally Lipschitz functions, it follows that they are all globally Lipschitz. ☐

C Proof of Theorem 4.1.

Let $\mathrm{conv}(A)$ denote the convex hull of A. First, we prove the following two propositions.

Proposition C.1. *Let e_i be the ith standard basis vector in $\mathbb{R}^n$. The Clarke's generalized gradient of $f(x) = x_{(p)}$ is given by*

$$\bar{\partial} x_{(p)} = \mathrm{conv}\{e_i \mid i \in E_{x_{(p)}}\}, \ \textit{where } E_{x_{(p)}} := \{i \mid x_i = x_{(p)}\}. \tag{67}$$

Proof. For any $x \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$, let $f^\circ(x; d)$ be the Clarke's generalized directional derivative at x along the direction d, i.e.,

$$f^\circ(x; d) := \limsup_{y \to x,\, t \to 0^+} \frac{f(y + td) - f(y)}{t}. \tag{68}$$

Define $d_{\max}(z) := \max\{d_i \mid i \in E_{z_{(p)}}\}$, $z \in \mathbb{R}^n$. First, we will show that

$$f^\circ(x; d) = d_{\max}(x). \tag{69}$$

Indeed, suppose that $L_{x_{(p)}} = \{i_1, \ldots, i_k\}$ and $E_{x_{(p)}} = \{i_{k+1}, \ldots, i_{k+l}\}$. Then, $k + 1 \le p \le k + l$. By the definitions in (66), there exists $\eta > 0$ such that for any $(y, t) \in B(x, \eta) \times (0, \eta)$ it holds that $(y + td)_i < (y + td)_j < (y + td)_k$ and $y_i < y_j < y_k$, for $\forall i \in L_{x_{(p)}}, \forall j \in E_{x_{(p)}}, \forall k \in G_{x_{(p)}}$. For any such (y, t), $y_{(p)} = (y_{i_{k+1}}, y_{i_{k+2}}, \ldots, y_{i_{k+l}})_{(p-k)}$ and $(y + td)_{(p)} = ((y + td)_{i_{k+1}}, (y + td)_{i_{k+2}}, \ldots, (y + td)_{i_{k+l}})_{(p-k)}$. Suppose without loss of generality that $y_{i_{k+1}} \le y_{i_{k+2}} \le \cdots \le y_{i_{k+l}}$. Then $y_{(p)} = y_{i_p}$. Let $j' \le p - k$ be the index such that $(y + td)_{i_{k+j'}} = \max\{(y + td)_{i_{k+1}}, (y + td)_{i_{k+2}}, \ldots, (y + td)_{i_p}\}$. Then $(y + td)_{(p)} = ((y + td)_{i_{k+1}}, (y + td)_{i_{k+2}}, \ldots, (y + td)_{i_{k+l}})_{(p-k)} \le \max\{(y + td)_{i_{k+1}}, (y + td)_{i_{k+2}}, \ldots, (y + td)_{i_p}\} = (y + td)_{i_{k+j'}}$. Furthermore, $j' \le p - k$ implies that $y_{i_p} \ge y_{i_{k+j'}}$. Therefore,

$$\frac{f(y + td) - f(y)}{t} = \frac{(y + td)_{(p)} - y_{i_p}}{t} \le \frac{(y + td)_{i_{k+j'}} - y_{i_{k+j'}}}{t} = d_{i_{k+j'}} \le d_{\max}(x). \tag{70}$$

Since (70) holds for any $(y, t) \in B(x, \eta) \times (0, \eta)$, it follows that

$$f^\circ(x; d) \le d_{\max}(x). \tag{71}$$

On the other hand, suppose $d_{i_{k+j^*}} = d_{\max}(x)$. Define $\zeta := 1 + \max\{|d_i| \mid i \in E_{x_{(p)}}\}$. There exists a sequence $y^{(m)} \to x$ as $m \to \infty$ such that for all m it holds that $y^{(m)}_{(p)} = (y^{(m)}_{i_{k+1}}, y^{(m)}_{i_{k+2}}, \ldots, y^{(m)}_{i_{k+l}})_{(p-k)} = y^{(m)}_{i_{k+j^*}}$ and $\min\{|y^{(m)}_{i_{k+a}} - y^{(m)}_{i_{k+b}}| \mid a \ne b, 1 \le a, b \le l\} = 2^{-m}\zeta$. Define $t^{(m)} := 2^{-m-2}$. Then

$$\frac{f(y^{(m)} + t^{(m)}d) - f(y^{(m)})}{t^{(m)}} = d_{i_{k+j^*}} = d_{\max}(x),\ \forall m. \tag{72}$$

Combining (72) with (71), we obtain (69).

Second, we will show that (67) holds. By definition, $\bar{\partial} f(x) := \{\xi \in \mathbb{R}^n \mid \xi^\top d \le f^\circ(x; d), \forall d \in \mathbb{R}^n\}$. On one hand, for $\forall \xi \in \operatorname{conv}\{e_i \mid i \in E_{x_{(p)}}\}$, ξ can be represented by $\xi = \sum_{i \in E_{x_{(p)}}} c_i e_i$, where $c_i \ge 0$ for all $i \in E_{x_{(p)}}$ and $\sum_{i \in E_{x_{(p)}}} c_i = 1$. Hence, $\xi^\top d \le d_{\max}(x) = f^\circ(x; d)$ for $\forall\, d \in \mathbb{R}^n$, which implies that $\operatorname{conv}\{e_i \mid i \in E_{x_{(p)}}\} \subseteq \bar{\partial} x_{(p)}$. On the other hand, for $\forall\, \xi \notin \operatorname{conv}\{e_i \mid i \in E_{x_{(p)}}\}$, it follows from separating hyperplane theorem that there exists $d \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $\xi^\top d > \alpha \ge \sup_{\mu \in \operatorname{conv}\{e_i \mid i \in E_{x_{(p)}}\}} \mu^\top d = d_{\max}(x) = f^\circ(x; d)$, which implies $\xi \notin \bar{\partial} x_{(p)}$. Therefore, $\bar{\partial} x_{(p)} \subseteq \operatorname{conv}\{e_i \mid i \in E_{x_{(p)}}\}$. Hence, (67) follows. $\square$

Proposition C.2. *For $\rho_{\text{Basel}} \in \{\rho_{\text{Basel2.5}}, \rho_{\text{Basel3}}\}$, there exists a closed and bounded set $\mathcal{C} \subset \mathbb{R}^n_+$ such that $0 \notin \mathcal{C}$ and $\bar{\partial}\rho_{\text{Basel}}(x) \subset \mathcal{C}$ for any $x \in \mathbb{R}^n$.*

Proof. Let $e_i^{[s]}$ be the ith standard basis in $\mathbb{R}^{n_s}$ and $E_{x_{(p)}^{[s]}} := \{1 \le i \le n_s \mid x_i^{[s]} = x_{(p)}^{[s]}\}$. By similar argument in the proof of Proposition C.1, it can be shown that

$$\bar{\partial} x_{(p)}^{[s]} = \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid i \in E_{x_{(p)}^{[s]}}\}. \tag{73}$$

Then by Theorem 2.3.3 and Theorem 2.3.10 in Clarke (1990), we have

$$\bar{\partial}\rho_{\text{Basel2.5}}(x) \subseteq \text{conv}\left(\cup_{i \in I_1(x)} \bar{\partial} f_i(x)\right) + \text{conv}\left(\cup_{i \in I_2(x)} \bar{\partial} f_i(x)\right), \tag{74}$$

where $f_1(x) = x_{(p_1)}^{[1]}$, and $\bar{\partial} f_1(x) = \text{conv}\{(e_i^{[1]}, 0, \ldots, 0) \mid i \in E_{x_{(p_1)}^{[1]}}\}$; $f_2(x) = \frac{k}{m_1}\sum_{s=1}^{m_1} x_{(p_s)}^{[s]}$, and $\bar{\partial} f_2(x) \subseteq \frac{k}{m_1}\sum_{s=1}^{m_1} \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid i \in E_{x_{(p_s)}^{[s]}}\}$; $I_1(x) := \{i \mid \max\{f_1(x), f_2(x)\} = f_i(x),\ i \in \{1, 2\}\}$; $f_3(x) = x_{(p_{m_1+1})}^{[m_1+1]}$, and $\bar{\partial} f_3(x) = \text{conv}\{(0, \ldots, 0, e_i^{[m_1+1]}, 0, \ldots, 0) \mid i \in E_{x_{(p_{m_1+1})}^{[m_1+1]}}\}$; $f_4(x) = \frac{l}{m_2}\sum_{s=m_1+1}^{m} x_{(p_s)}^{[s]}$, and $\bar{\partial} f_4(x) \subseteq \frac{l}{m_2}\sum_{s=m_1+1}^{m} \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid i \in E_{x_{(p_s)}^{[s]}}\}$; $I_2(x) := \{i \mid \max\{f_3(x), f_4(x)\} = f_i(x),\ i \in \{3, 4\}\}$. Define $A_1 := \text{conv}\{(e_i^{[1]}, 0, \ldots, 0) \mid 1 \le i \le n_1\}$, $A_2 := \frac{k}{m_1}\sum_{s=1}^{m_1} \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid 1 \le i \le n_s\}$, $A_3 := \text{conv}\{(0, \ldots, 0, e_i^{[m_1+1]}, 0, \ldots, 0) \mid 1 \le i \le n_{m_1+1}\}$, $A_4 := \frac{l}{m_2}\sum_{s=m_1+1}^{m} \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid 1 \le i \le n_s\}$, and $\mathcal{C} := \text{conv}(A_1 \cup A_2) + \text{conv}(A_3 \cup A_4)$. Then, $\mathcal{C}$ is compact and $0 \notin \mathcal{C}$; in addition, it follows from (74) that $\bar{\partial}\rho_{\text{Basel2.5}}(x) \subset \mathcal{C}$ for any $x \in \mathbb{R}^n$.

By (14) and (73), we have

$$\begin{aligned}\bar{\partial}\rho_{\text{CVaR}_\alpha}(x^{[s]}) \subseteq{}& \frac{p_s - \alpha n_s}{(1-\alpha)n_s}\text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid i \in E_{x_{(p_s)}^{[s]}}\} \\ &+ \frac{1}{(1-\alpha)n_s}\sum_{j=p_s+1}^{n_s} \text{conv}\{(0, \ldots, 0, e_i^{[s]}, 0, \ldots, 0) \mid i \in E_{x_{(j)}^{[s]}}\}.\end{aligned}$$

Then, we can show by similar argument that the conclusion also holds for $\rho_{\text{Basel}} = \rho_{\text{Basel3}}$. ∎

The proof of Theorem 4.1 is as follows.

Proof of Theorem 4.1. (i) Since $\rho(x)$ is locally Lipschitz and $\rho_{\text{Basel}}(x)$ is globally Lipschitz on $\mathbb{R}^n$ (by Lemma 4.1), it follows from Proposition 2.1.2 in Clarke (1990) that $\bar{\partial}\rho(x)$ and $\bar{\partial}\rho_{\text{Basel}}(x)$ exist on $\mathbb{R}^n$. Then the first part of the theorem follows from the corollary of Proposition 2.4.3 and Theorem 2.3.10 in Clarke (1990).

(ii) To prove part (ii), we first show that

$$\lim_{j\to\infty} x^{(j+1)} - x^{(j)} = 0,\ \lim_{j\to\infty} y^{(j+1)} - y^{(j)} = 0, \text{ and } \lim_{j\to\infty} u^{(j+1)} - u^{(j)} = 0. \tag{75}$$

Since $\mathcal{U}_{r_0}$ is closed and bounded, the sequence $\{u^{(j)}\}$ is bounded. It follows from (26) and (27) that $x^{(j+1)} = (\lambda^{(j+1)} - \lambda^{(j)})/(\beta_1\sigma_1) - \tilde{R}u^{(j+1)}$ and $y^{(j+1)} = (\pi^{(j+1)} - \pi^{(j)})/(\beta_2\sigma_2) - \tilde{Y}u^{(j+1)}$, which in combination with boundedness of $\{u^{(j)}\}$ and assumed boundedness of $\{(\lambda^{(j)}, \pi^{(j)})\}$ implies that $\{(x^{(j)}, y^{(j)})\}$ is bounded. Thus, $\{(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)})\}$ is bounded, and then the continuity of the augmented Lagrangian function (22) implies that $\{\mathcal{L}(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)})\}$ is bounded.

Note that the augmented Lagrangian function $\mathcal{L}$ is strongly convex with respect to the variable u. Therefore, it holds that for any u and Δu,

$$\mathcal{L}(x, y, u + \Delta u, \lambda, \pi) - \mathcal{L}(x, y, u, \lambda, \pi) \geq \partial_u \mathcal{L}(x, y, u, \lambda, \pi)^\top \Delta u + c\|\Delta u\|^2, \tag{76}$$

where $c > 0$ is constant. In addition, since $u^{(j+1)}$ minimizes (25) and $u^{(j)} \in \mathcal{U}_{r_0}$, it follows that

$$\partial_u \mathcal{L}(x^{(j+1)}, y^{(j+1)}, u^{(j+1)}, \lambda^{(j)}, \pi^{(j)})^\top (u^{(j)} - u^{(j+1)}) \geq 0. \tag{77}$$

Combining (76) and (77), we obtain

$$\mathcal{L}(x^{(j+1)}, y^{(j+1)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)}) - \mathcal{L}(x^{(j+1)}, y^{(j+1)}, u^{(j+1)}, \lambda^{(j)}, \pi^{(j)}) \geq c\|u^{(j+1)} - u^{(j)}\|^2. \tag{78}$$

Since $x^{(j+1)}$ minimizes (23) and $y^{(j+1)}$ minimizes (24), it follows from (26), (27), and (78) that

$$\begin{aligned} &\mathcal{L}(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)}) - \mathcal{L}(x^{(j+1)}, y^{(j+1)}, u^{(j+1)}, \lambda^{(j+1)}, \pi^{(j+1)}) \\ &+ \frac{1}{\beta_1\sigma_1}\|\lambda^{(j)} - \lambda^{(j+1)}\|^2 + \frac{1}{\beta_2\sigma_2}\|\pi^{(j)} - \pi^{(j+1)}\|^2 \geq c\|u^{(j+1)} - u^{(j)}\|^2. \end{aligned} \tag{79}$$

Since $\sum_{j=1}^{\infty}(\|\lambda^{(j+1)} - \lambda^{(j)}\|^2 + \|\pi^{(j+1)} - \pi^{(j)}\|^2) < \infty$ and $\{\mathcal{L}(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)})\}$ is bounded, it follows from (79) that

$$\sum_{j=1}^{\infty} \|u^{(j+1)} - u^{(j)}\|^2 < \infty, \tag{80}$$

which implies that

$$\lim_{j\to\infty} u^{(j+1)} - u^{(j)} = 0. \tag{81}$$

Since $\sum_{j=1}^{\infty}(\|\lambda^{(j+1)} - \lambda^{(j)}\|^2 + \|\pi^{(j+1)} - \pi^{(j)}\|^2) < \infty$, it follows that $\lim_{j\to\infty} \lambda^{(j+1)} - \lambda^{(j)} = 0$, which in combination with (26) implies that

$$\lim_{j\to\infty} x^{(j+1)} + \tilde{R}u^{(j+1)} = 0. \tag{82}$$

By (81) and (82), we obtain $\lim_{j\to\infty} x^{(j+1)} - x^{(j)} = 0$. By similar argument, we obtain $\lim_{j\to\infty} y^{(j+1)} - y^{(j)} = 0$.

For any limit point $\bar{u}$ of the sequence $\{u^{(j)}\}$, there exists a subsequence $u^{(k_i)} \to \bar{u}$ as $i \to \infty$. Since $\{(x^{(j)}, y^{(j)}, u^{(j)}, \lambda^{(j)}, \pi^{(j)})\}$ is bounded, there exists a further subsequence

$\{j_i\} \subseteq \{k_i\}$ such that $(x^{(j_i)}, y^{(j_i)}, u^{(j_i)}, \lambda^{(j_i)}, \pi^{(j_i)}) \to (\bar{x}, \bar{y}, \bar{u}, \bar{\lambda}, \bar{\pi})$ as $i \to \infty$. Clearly, we obtain from (81) and (82) that

$$\bar{x} + \tilde{R}\bar{u} = \lim_{i\to\infty} x^{(j_i)} + \tilde{R}u^{(j_i)} = \lim_{i\to\infty} x^{(j_i)} + \tilde{R}u^{(j_i-1)} = 0. \tag{83}$$

A similar argument leads to $\bar{y} + \tilde{Y}\bar{u} = 0$.

The first-order optimality condition of (23) in the j_ith iteration is

$$0 \in \sigma_1(x^{(j_i)} + \tilde{R}u^{(j_i-1)}) + \lambda^{(j_i-1)} + \eta^{(j_i)}\bar{\partial}\rho_{\text{Basel}}(x^{(j_i)}), \tag{84}$$

$$\eta^{(j_i)}(\rho_{\text{Basel}}(x^{(j_i)}) - C_0) = 0, \tag{85}$$

for some $\eta^{(j_i)} \geq 0$. Since $\{\bar{\partial}\rho_{\text{Basel}}(x^{(j_i)})\}$ is bounded away from zero (by Proposition C.2) and $\{x^{(j_i)}\}, \{u^{(j_i)}\}$ and $\{\lambda^{(j_i)}\}$ are bounded, it follows from (84) that the sequence $\{\eta^{(j_i)}\}$ is bounded. Hence, $\{\eta^{(j_i)}\}$ has a subsequence that converges. For the sake of simplification of notation, we still denote the subsequence as $\{\eta^{(j_i)}\}$ and denote η as its limit. Since $\lim_{j_i\to\infty} \lambda^{(j_i)} - \lambda^{(j_i-1)} = 0$, it follows that $\lim_{i\to\infty} \lambda^{(j_i-1)} = \bar{\lambda}$. Then, applying Proposition 2.1.5 in Clarke (1990) and noting the uniform boundedness of $\bar{\partial}\rho_{\text{Basel}}(x^{(j_i)})$ (by Proposition C.2) and (83), we obtain from (84) and (85) that

$$\bar{\lambda} \in -\eta\bar{\partial}\rho_{\text{Basel}}(\bar{x}), \tag{86}$$

$$\eta(\rho_{\text{Basel}}(\bar{x}) - C_0) = 0. \tag{87}$$

The first-order optimality condition of (24) in the j_ith iteration is

$$0 \in \bar{\partial}\rho(y^{(j_i)}) + \pi^{(j_i-1)} + \sigma_2(y^{(j_i)} + \tilde{Y}u^{(j_i-1)}). \tag{88}$$

Applying Proposition 2.1.5 in Clarke (1990), and taking limit on both sides of (88), we obtain

$$\bar{\pi} \in -\bar{\partial}\rho(\bar{y}). \tag{89}$$

The first-order optimality condition of (25) in the j_ith iteration leads to

$$\tilde{R}^\top\lambda^{(j_i-1)} + \sigma_1\tilde{R}^\top(x^{(j_i)} + \tilde{R}u^{(j_i)}) + \tilde{Y}^\top\pi^{(j_i-1)} + \sigma_2\tilde{Y}^\top(y^{(j_i)} + \tilde{Y}u^{(j_i)}) + \zeta^{(j_i)} = 0, \tag{90}$$

where $\zeta^{(j_i)} \in \mathcal{N}_{\mathcal{U}_{r_0}}(u^{(j_i)})$, which is the normal cone to $\mathcal{U}_{r_0}$ at $u^{(j_i)}$. It follows from (90) and the convergence of $\{(x^{(j_i)}, y^{(j_i)}, u^{(j_i)}, \lambda^{(j_i)}, \pi^{(j_i)})\}$ that $\bar{\zeta} := \lim_{i\to\infty} \zeta^{(j_i)}$ is well defined. Since $\mathcal{U}_{r_0}$ is compact and convex, it follows from Proposition 2.4.4 in Clarke (1990) that the normal cone $\mathcal{N}_{\mathcal{U}_{r_0}}(u)$ coincides with the cone of normals. Applying Proposition 2.1.5 in Clarke (1990) to the cone of normals, we obtain $\bar{\zeta} \in \mathcal{N}_{\mathcal{U}_{r_0}}(\bar{u})$. Taking limit on both sides of (90) and applying (86), (87), (89) and $\bar{\zeta} \in \mathcal{N}_{\mathcal{U}_{r_0}}(\bar{u})$, we obtain (38) with u being $\bar{u}$. This completes the proof. □

References

C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26:1487–1503, 2002.

T. Adrian and M. K. Brunnermeier. CoVaR. Federal Reserve Bank of New York Staff Reports, 348, 2008.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

X. D. Bai, J. Sun, X. J. Zheng, and X. L. Sun. Penalty decomposition methods for probabilistically constrained convex programs. Working paper, Fudan University, 2012. Available at: http://www.optimization-online.org/DB_HTML/2012/04/3448.html.

Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards: A revised framework. Document, Bank for International Settlements, Basel, Switzerland, 2006.

Basel Committee on Banking Supervision. Revisions to the Basel II market risk framework. Document, Bank for International Settlements, Basel, Switzerland, 2009.

Basel Committee on Banking Supervision. Basel III: A global regulatory framework for more resilient banks and banking systems. Document, Bank for International Settlements, Basel, Switzerland, December 2010.

Basel Committee on Banking Supervision. Fundamental review of the trading book. Consultative Document, Bank for International Settlements, Basel, Switzerland, May 2012.

Board of Governors of the Federal Reserve Systems. Enhanced prudential standards and early remediation requirements for covered companies. *Federal Register*, 77(3):594–663, January 2012.

A. Chairawongse, S. Kiatsupaibul, S. Tirapat, and B. Van Roy. Portfolio selection with qualitative input. *Journal of Banking and Finance*, 36(2):489–496, 2012.

F. H. Clarke. *Optimization and Nonsmooth Analysis*. SIAM, Philadelphia, PA, 1990.

X. T. Cui, S. S. Zhu, X. L. Sun, and D. Li. Nonlinear portfolio selection using approximate parametric Value-at-Risk. *Journal of Banking and Finance*, 37:2124–2139, 2013.

Q. Dai, K. J. Singleton, and W. Yang. Regime shifts in a dynamic term structure model of U.S. treasury bonds. *Review of Financial Studies*, 20:1669–1706, 2007.

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

J. D. Hamilton. What's real about the business cycle? *Federal Reserve Bank of St. Louis Review*, 87:435–452, 2005.

B. He and X. Yuan. On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method. *SIAM Journal on Numerical Analysis*, 50(2):700–709, 2012.

M. Hong and Z.-Q. Luo. On the linear convergence of the alternating direction method of multipliers. *ArXiv e-prints*, 2012.

G. Iyengar and K. C. Ma. Fast gradient descent method for mean-CVaR optimization. *Annals of Operations Research*, 205:203–212, 2013.

P. Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, Boston, 3 edition, 2007.

N. E. Karoui, A. E. B. Lim, and G. Y. Vahn. Performance based regularization in mean-CVaR portfolio optimization. Working paper, University of California, Berkeley, 2011.

S. Kou. A jump diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

A. E. B. Lim, J. G. Shanthikumar, and G.-Y. Vahn. Conditional Value-at-Risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163 – 171, 2011. ISSN 0167-6377.

R. T. Rockafellar and S. Uryasev. Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance*, 26:1443–1471, 2002.

C. A. Sims and T. Zha. Were there regime switches in U.S. monetary policy? *American Economic Review*, 96:54–81, 2006.

S. S. Wang, V. R. Young, and H. H. Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, 21:173–183, 1997.

Z. Wen, D. Goldfarb, and W. Yin. Alternating direction augmented Lagrangian methods for semidefinite programming. *Mathematical Programming Computation*, 2:203–230, 2010.