Quantitative Finance, 2021
Vol. 21, No. 10, 1707–1721, https://doi.org/10.1080/14697688.2021.1879392

Robust portfolio rebalancing with cardinality and diversification constraints

ZHIHUA ZHAO†, FENGMIN XU*†, DONGLEI DU‡ and WANG MEIHUA§

†School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710061, People's Republic of China
‡Faculty of Business Administration, University of New Brunswick, Fredericton, Canada
§School of Economics and Management, Xidian University, Xi'an 710071, People's Republic of China

(*Received 13 April 2019; accepted 15 January 2021; published online 26 March 2021*)

In this paper, we develop a robust conditional value at risk (CVaR) optimal portfolio rebalancing model under various financial constraints to construct sparse and diversified rebalancing portfolios. Our model includes transaction costs and double cardinality constraints in order to capture the trade-off between the limit of investment scale and the diversified industry coverage requirement. We first derive a closed-form solution for the robust CVaR portfolio rebalancing model with only transaction costs. This allows us to conduct an industry risk analysis for sparse portfolio rebalancing in the absence of diversification constraints. Then, we attempt to remedy the hidden industry risk by establishing a new robust portfolio rebalancing model with both sparse and diversified constraints. This is followed by the development of a distributed-version of the Alternating Direction Method of Multipliers (ADMM) algorithm, where each subproblem admits a closed-form solution. Finally, we conduct empirical tests to compare our proposed strategy with the standard sparse rebalancing and no-rebalancing strategies. The computational results demonstrate that our rebalancing approach produces sparse and diversified portfolios with better industry coverage. Additionally, to measure out-of-sample performance, two superiority indices are created based on worst-case CVaR and annualized return, respectively. Our ADMM strategy outperforms the sparse rebalancing and no-rebalancing strategies in terms of these indices.

Keywords: Portfolio rebalancing; Cardinality constraint; Diversification constraint; Sparse projection; ADMM.

JEL Classifications: G11, C61

1. Introduction

As a risk management strategy, portfolio rebalancing aims to decrease investor risk exposure by periodically and efficiently buying or selling assets. In this paper, we focus on two core issues of portfolio rebalancing: risk measures and rebalancing philosophy. We consider a frictional market over a given investment horizon responsive to cash changes and transaction costs. We attempt to balance between the limits of investment scale and a diversified industry coverage.

Risk measures are quantified either as the deviation between the constructed portfolio and a benchmark index, or as an assessment of the potential for absolute loss of the portfolio (Rockafellar *et al.* 2006). The former type of risk measure has long dominated much of portfolio analysis. It relies on historical performance of securities to collect quality 'master funds' through summarizing the corresponding tendencies in a financial market (Markowitz 1970). This approach can be further divided into either value-based formulations by minimizing the mean-absolute deviation (Sharpe 1971, Konno and Yamazaki 1991, Fang *et al.* 2005), or return-based ones by minimizing an utility function, referred to as tracking error (Takeda *et al.* 2013, SantAnna *et al.* 2017), semi-absolute deviation (Speranza 1993, Chiodi *et al.* 2003), and skewness and kurtosis (Yu and Lee 2011), etc. (see Strub and Baumann 2018, for a detailed comparison between value-based and return-based portfolios).

The latter type of risk measure, on the other hand, has involved a series of extensions based on the concept of Value-at-Risk (VaR) (Basak and Shapiro 2001, Gaivoronski and Pflug 2005). Rockafellar and Uryasev (2000) and Rockafellar and Uryasev (2002) proposed the Conditional Value-at-Risk (CVaR). CVaR is defined as the conditional expectation of losses in top $100(1-\beta)\%$ over a given investment horizon (e.g. $\beta = 0.95, 0.99$). CVaR is more commonly adopted than

*Corresponding author. Email: fengminxu@mail.xjtu.edu.cn

VaR in practice because of some better properties. For example, CVaR is a coherent risk measure (Artzner *et al.* 1999), and minimizing CVaR can often lead to the formulation of tractable optimization problems (e.g. convex programs, even linear programs) (Rockafellar and Uryasev 2002, Zhu and Fukushima 2009). However, the implementation of VaR, CVaR and any other distribution-based risk measure is frequently plagued by the unavailability of a statistical distribution of market parameters. These market parameters are often estimations from sample data, inevitably resulting in statistical and modeling errors. In practice, many efficient approaches have been proposed for mitigating this sensitivity in the estimates of the relevant market parameters. One such approach is to directly analyze the CVaR utility function and its partial derivative functions with respect to the related parameters, e.g. CVaR sensitivity estimation (Hong and Liu 2009, He 2019). Some other approaches mentioned in the literature are embedding a variables-reparametrized optimization into the VaR (or CVaR) risk measure, e.g. the VaR-based Black–Litterman model (Lejeune 2011), the Bayesian shrinkage estimation methods (Ledoit and Wolf 2003, 2004), and the robust VaR (or CVaR) risk measure (Ghaoui *et al.* 2003, Zhu and Fukushima 2009).

Robust risk measures have been developed, aiming to optimize asset allocation under the worst-case risk analysis over uncertainty sets of candidate distributions. Lobo and Boyd (2000) considered a robust portfolio allocation problem in the classical mean-variance (MV) framework (Markowitz 1952), where the means and covariances of asset returns are assumed to belong to an uncertainty set (box or ellipsoid). This was the first attempt to include second moment uncertainty. Ghaoui *et al.* (2003) defined the robust VaR over various uncertain moment sets. When the first and second moments are known, a closed-form solution for minimizing worst-case VaR is obtained; otherwise, the optimization problem can be cast as a semidefinite program (SDP). They also conducted further analysis for various other partial moment information of the distribution. Zhu and Fukushima (2009) proposed robust CVaR, and proved that it is still a coherent risk measure. They then investigated the minimization of the robust CVaR over three different uncertainty sets and applied these three scenarios to robust portfolio problems, formulated as tractable linear programs or second-order cone programs (SOCP). Given first and second moments, tight upper bounds for some special utility functions have been established. For example, Natarajan and Linyi (2007) established results for linear utility functions, and Natarajan *et al.* (2010) established results for piecewise-linear concave utility functions. Natarajan *et al.* (2010) also derived a closed-form solution for the worst-case CVaR without any constraints. Other generalizations are further discussed in Natarajan *et al.* (2010), Chen *et al.* (2011) and Fabozzi *et al.* (2010).

Rebalancing can be achieved either by market timing or by selecting a fixed time-point (FTP). Here, we only pay attention to a class of very popular FTP rebalancing strategies (a.k.a., sparsity-induced rebalancing strategies), which are directly relevant to our work. These are normally designed to either require risk measurement with sparsity-induced regularization (penalty) with a 'tunable' coefficient, or to optimize a certain utility function subject to sparse constraints imposed on the portfolio holdings. These penalty terms and sparse constraints generally involve some norms, such as the sum of the absolute values of the portfolio weights (the l_1 norm) (Tibshirani 1996), or the cardinality of the portfolio holdings (the l_0 norm) (Blumensath and Davies 2009). These norms can facilitate the treatment of trading costs, e.g. the l_1 norm is interpreted as a fixed bid-ask spread which is a principal cost for institutional investors, while the l_0 norm is considered to be volume-independent 'overhead' cost which is the main cost for retail investors (see Brodie *et al.* 2009, for a detailed explanation).

For the sparsity-constrained approach, Chang *et al.* (2000) introduced cardinality constraints into the classical MV model and illustrated the corresponding change in the shape of efficient frontier. Beasley *et al.* (2003) proposed an evolutionary heuristic algorithm for the cardinality-constrained index tracking problem. Shaw *et al.* (2008) developed a dedicated Lagrangian relaxation method for cardinality-constrained portfolio selection. Woodside-Oriakhi *et al.* (2011) considered an MV model with buy-in thresholds and cardinality constraints and applied a heuristic algorithm to find the sparse efficient frontier. Woodside-Oriakhi *et al.* (2013) theoretically indicated the impact of the investment horizon on the portfolio variance and return for a portfolio rebalancing model with a linear transaction cost, and then experimentally explored the change in the efficient frontier. Gao and Li (2013) examined geometric properties of cardinality-constrained MV portfolio selection. By designing relaxed problems under the circumscribed box, ball and axis-aligned ellipsoid uncertain sets, they obtained the resulting efficient lower bounds. They then proposed a dedicated branch-and-bound algorithm, which integrated these lower bounding results. Kyrillidis *et al.* (2013) derived efficient sparse projection over the simplex, and then applied it to portfolio optimization. Xu *et al.* (2016) and Xu *et al.* (2019) generalized the bounds imposed on portfolio weights of Kyrillidis *et al.* (2013), and established more general closed-form expressions.

For the regularization approach, Brodie *et al.* (2009) introduced an l_1 penalty into the classical MV model to encourage sparse portfolios and to account for transaction costs. DeMiguel *et al.* (2007) and DeMiguel *et al.* (2009) discussed several regularization techniques for portfolio construction, including the imposition of constraints on l_p ($p = 1, 2$) norms of portfolio weights. Takeda *et al.* (2013) proposed an l_2 regularized model with l_0 norm constraint for index tracking problem. Fan *et al.* (2008) provided mathematical insights into utility approximation with the gross-exposure constraint. Chen *et al.* (2013) proposed an interior point algorithm to obtain an approximate second-order KKT solution for the l_p $(0 < p < 1)$ norm regularized portfolio selection model. Xu *et al.* (2015) established an $l_{1/2}$ regularized index tracking model, and proposed an efficient hybrid half thresholding algorithm (see Bruder *et al.* 2013, for more regularization techniques). Moreover, discussions about the impact of the tunable coefficient on the portfolio's performance can be found in Schwarz (1978), Tibshirani (1996), Chen *et al.* (2010) and Chen *et al.* (2014).

However, we note that the sparse setting can only restrict the amount of assets in the entire portfolio, not the industry distribution to which the assets belong. Therefore, this kind of strategy may result in an over-concentrated portfolio in a few industries or sectors,† with the potential for loss due to sudden negative shocks to these sectors or industries. On the other hand, adding cardinality and/or diversification constraints generally leads to NP-hard problems (Chen *et al.* 2014), for which there exists no polynomial time algorithm for global solution. To our best knowledge, Bonami and Lejeune (2009) made the first meaningful attempt to include diversification constraints into the MV portfolio selection problem, and then proposed a nonlinear branch-and-bound algorithm to solve it. Branda (2013) and Branda (2015) established a Data Envelopment Analysis (DEA) portfolio model with diversification and cardinality constraints, but no method was proposed to solve this model.

In this paper, we propose a new sparse and diversified portfolio rebalancing model, where we minimize the worst-case CVaR and the penalty term of transaction cost subject to two cardinality constraints that balance the limits of investment scale and industry coverage. We represent evolving market environments via customized 'diversification constraints', and cash changes via 'budget constraints'. We also include 'buy-in threshold constraints', which are box constraints over the support set of the portfolio. For the solution method, we first derive a closed-form solution for a robust CVaR rebalancing model with transaction costs. We then use this formula to perform a case study to illustrate the industry risk of sparse rebalancing in the absence of diversification constraints. Then, we attempt to remedy the hidden industry risk by establishing a new sparse and diversified portfolio rebalancing model. This is followed by the development of a distributed consensus-type Alternating Direction Method of Multipliers (ADMM), where a central sum variable is introduced to break the original problem down into three tractable subproblems. At each iteration, we only need to solve alternately three subproblems with closed-form solutions. Finally, we conduct an out-of-sample performance analysis of the proposed model and method with real market data sets from the Chinese stock market.

The contribution of this paper is twofold. From the perspective of the optimal rebalancing of a portfolio: (i) we conduct industry risk analysis for the sparse CVaR rebalancing strategy quantitatively to reveal that it may lead to an over-concentrated portfolio in few industries.‡ (ii) To address this issue, we propose a sparse and diversified rebalancing strategy that balances diversification across industries and sparseness of investment scale. The adjusted portfolio of our strategy not only inherits the advantages of sparse rebalancing for easy management, but also increases its ability to resist industry shocks. From the perspective of the solution of our proposed model (the essence is an SOCP with 'double' cardinality constraints): (i) we develop a distributed consensus-type ADMM, in which each subproblem admits a closed-form solution. The out-of-sample performance in actual data sets verifies the effectiveness of our model and algorithm. (ii) Moreover, the sparse projection operator established for the subproblems is of independent interest and may find applications to other sparse models.

In the remainder of this section, we introduce the notation used in the paper. In Section 2, we derive a closed-form solution for a robust CVaR rebalancing model. Then we conduct industry risk analysis for sparse rebalancing in Section 3. In Section 4, we describe the new portfolio rebalancing model and propose an efficient ADMM for the proposed model. In Section 5, we conduct empirical tests to compare our strategy with two well-established strategies for portfolio rebalancing. Finally, we draw concluding remarks in Section 6.

1.1. Notations

Let $\mathbb{Z}^N$ and $\mathbb{R}^N$ denote the N-dimensional integer set and Euclidean space, respectively. The symbols $\mathbb{Z}^N_+$ and $\mathbb{R}^N_+$ stand for the set of nonnegative number of $\mathbb{Z}^N$ and $\mathbb{R}^N$, respectively. For any $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^N$, denote the Hadmard product $\boldsymbol{x}\circ\boldsymbol{y}=(x_1y_1,\dots,x_ny_n)^T$, and $\boldsymbol{x}\succeq(\succ)\boldsymbol{y}$ if $x_j\geq(>)y_j$ for any $j\in\{1,2,\dots,N\}$. For any $a,b,c\in\mathbb{R}$ $(a<b)$, denote $a/\boldsymbol{x}=(a/x_1,a/x_2,\dots,a/x_N)^T$. Let $\mathrm{In}_{[a,b]}(c)$ denote an indicative function over the interval $[a,b]$; that is, $\mathrm{In}_{[a,b]}(c)=1$ if $c\in[a,b]$; and 0, otherwise. Define $\Pi_{[a,b]}(\boldsymbol{x})$ as the vector of the projection of $\boldsymbol{x}$ over the interval $[a,b]$, namely,

$$[\Pi_{[a,b]}(\boldsymbol{x})]_i=\begin{cases}a, & \text{if } x_i\leq a;\\ x_i, & \text{if } a<x_i<b;\\ b, & \text{if } x_i\geq b.\end{cases}\qquad \forall i=1,2,\dots,N.$$

For any $s\in\mathbb{Z}_+$, denote $I_s(\boldsymbol{x})$ as the indices of the s largest entries of $\boldsymbol{x}$. Define $\mathbf{1}$ and $\boldsymbol{I}$ as the vector of all ones and the identity matrix, respectively, of which the dimension can be known from the context. The set of all $M\times N$ matrices with real entries is denoted by $\mathbb{R}^{M\times N}$. For any $X\in\mathbb{R}^{M\times N}$, $X^\dagger$ denotes the Moore–Penrose pseudo inverse of X. The spaces of $N\times N$ diagonal and symmetric matrices are denoted by $\mathbb{D}^n$ and $\mathbb{S}^n$, respectively. If $X\in\mathbb{S}^N$ is positive semidefinite (positive definite), we write $X\succeq 0$ $(X\succ 0)$. The cone of positive semidefinite matrices is denoted by $\mathbb{S}^N_+$, and similarly with $\mathbb{D}^N_+$. For any $X\in\mathbb{S}^N_+$, $\sigma(X)$ denotes the N-dimensional vector consisting of eigenvalues of X. We define the operator $\mathcal{D}:\mathbb{R}^N\mapsto\Re^{N\times N}$ as follows:

$$\mathcal{D}_{i,j}(\boldsymbol{x})=\begin{cases}x_i, & \text{if } i=j;\\ 0, & \text{otherwise.}\end{cases}\qquad \forall\,\boldsymbol{x}\in\mathbb{R}^N.$$

2. Closed-form solution for portfolio rebalancing model

In this section, we establish a closed-form solution for a robust CVaR model with transaction costs for portfolio rebalancing, which will be called upon later.

We consider N securities with random return vector by $\boldsymbol{r}=(r_1,\dots,r_N)^T$. Let $\mathbb{E}(\boldsymbol{r})=\mu$ be mean vector of the returns, and let $Cov(\boldsymbol{r})=\mathbb{E}[(\boldsymbol{r}-\mu)(\boldsymbol{r}-\mu)^T]=\Sigma$ be the covariance matrix of returns. A portfolio is defined as weights

† Thereafter, unless explicitly stated, the 'few' in the text emphasizes that the number of economic industries linked with the adjusted portfolio is very small. For instance, in the financial case in Section 3, the optimal portfolio holdings after sparse rebalancing are concentrated in 1 or 2 industries and we say they are over-concentrated in a few industries.

‡ The corresponding cases can refer to the financial example in Section 3, and the experimental results of the down-up1 subset in Section 5.

w_i $(i = 1, \ldots, N)$, that represent the amount of capital to be invested in each security. The initial capital S_0 is fully invested, which implies that the initial portfolio $\boldsymbol{w}_0 \in \mathbb{R}^N$ satisfies $e^T \boldsymbol{w}_0 = S_0$. Denote $\Delta \boldsymbol{w} \subseteq \mathbb{R}^N$ and $\Delta S \in \mathbb{R}$ as the traded amounts of N securities and the external capital flow, respectively, where $\Delta S > 0$ means investment, while $\Delta S < 0$ means divestment. We consider linear transaction costs with the proportion vector $\boldsymbol{c} \in \mathbb{R}_+^N$ for the securities. Suppose that $\boldsymbol{r}$ follows a continuous distribution with density function $p(\cdot)$. Let $f(\Delta \boldsymbol{w}, \boldsymbol{r})$ be the loss function of portfolio rebalancing, satisfying $\mathbb{E}[|f(\Delta \boldsymbol{w}, \boldsymbol{r})|] < +\infty$, so that CVaR and worst-case CVaR are well-defined.

For any fixed $\Delta \boldsymbol{w}$, we represent the probability of $f(\Delta \boldsymbol{w}, \boldsymbol{r})$ no more than a threshold α as

$$\psi(\Delta \boldsymbol{w}, \alpha) = \mathbb{P}\{f(\Delta \boldsymbol{w}, \boldsymbol{r}) \leq \alpha\} = \int_{f(\Delta \boldsymbol{w}, \boldsymbol{r}) \leq \alpha} p(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.$$

Then, the VaR of $\Delta \boldsymbol{w}$ with a confidence level β can be defined as

$$\mathrm{VaR}_\beta(\Delta \boldsymbol{w}) = \min\{\alpha \in \mathbb{R} : \psi(\Delta \boldsymbol{w}, \alpha) \geq \beta\}.$$

We can thereby formulate mathematically the corresponding CVaR as

$$\begin{aligned}\mathrm{CVaR}_\beta(\Delta \boldsymbol{w}) &= \mathbb{E}\left[f(\Delta \boldsymbol{w}, \boldsymbol{r}) | f(\Delta \boldsymbol{w}, \boldsymbol{r}) \geq \mathrm{VaR}_\beta(\Delta \boldsymbol{w})\right] \\ &= \frac{1}{1-\beta} \int_{f(\Delta \boldsymbol{w}, \boldsymbol{r}) \geq \mathrm{VaR}_\beta(\Delta \boldsymbol{w})} f(\Delta \boldsymbol{w}, \boldsymbol{r}) p(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.\end{aligned}$$

Further, we can also calculate CVaR by minimizing the resulting auxiliary function $G_\beta(\Delta \boldsymbol{w}, \alpha)$ with respect to the variable $\alpha \in \mathbb{R}$ (Rockafellar and Uryasev 2000, 2002), where

$$\begin{aligned}G_\beta(\Delta \boldsymbol{w}, \alpha) &= \alpha + \frac{1}{1-\beta} \mathbb{E}\left[(f(\Delta \boldsymbol{w}, \boldsymbol{r}) - \alpha)_+\right] \\ &= \alpha + \frac{1}{1-\beta} \int_{\boldsymbol{r} \in \mathbb{R}^N} (f(\Delta \boldsymbol{w}, \boldsymbol{r}) - \alpha)_+ p(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.\end{aligned}$$

Thus, we have

$$\mathrm{CVaR}_\beta(\Delta \boldsymbol{w}) = \min_{\alpha \in \mathbb{R}} G_\beta(\Delta \boldsymbol{w}, \alpha).$$

Suppose that the uncertainty set $\mathcal{P}$ of the density function $p(\cdot)$ is defined by

$$\mathcal{P} = \{p(\boldsymbol{r}) : \mathbb{E}(\boldsymbol{r}) = \mu, Cov(\boldsymbol{r}) = \Sigma \succ 0\}.$$

We then define the worst-case CVaR for fixed $\boldsymbol{w} \in \mathbb{R}^N$ over the distribution set $\mathcal{P}$ as

$$\begin{aligned}\mathrm{WCVaR}_\beta(\Delta \boldsymbol{w}) &= \sup_{p(\boldsymbol{r}) \in \mathcal{P}} \mathrm{CVaR}_\beta(\Delta \boldsymbol{w}) \\ &= \sup_{p(\boldsymbol{r}) \in \mathcal{P}} \min_{\alpha \in \mathbb{R}} G_\beta(\Delta \boldsymbol{w}, \alpha).\end{aligned}$$

Since $G_\beta(\Delta \boldsymbol{w}, \alpha)$ is a convex function of α (Chen *et al.* 2011), it follows from the max-min theorem (see Chapter 6 of Bazaraa *et al.* 2013) that

$$\mathrm{WCVaR}_\beta(\Delta \boldsymbol{w}) = \min_{\alpha \in \mathbb{R}} \sup_{p(\boldsymbol{r}) \in \mathcal{P}} G_\beta(\Delta \boldsymbol{w}, \alpha).$$

Given a fixed rebalancing time t, the optimal rebalancing amount Δw_i for asset i is decided at time t^- and completed at time t^+. We set the transient return of asset i at time t to be $r_{i,t}$. Then, a transient effect on the transaction cost of every asset i can be formulated as

$$\mathrm{TC}(\Delta w_i) = c_i(1 + r_{i,t})|\Delta w_i|.$$

Now, we are ready to establish a robust CVaR rebalancing model with budget constraint, denoted as $\Omega_1 = \{\Delta \boldsymbol{w} \in \mathbb{R}^N : \boldsymbol{1}^T \Delta \boldsymbol{w} = \Delta S\}$, as follows:

$$\min_{\Delta \boldsymbol{w} \in \Omega_1} \quad \kappa \mathrm{WCVaR}_\beta(\Delta \boldsymbol{w}) - \mathbb{E}(\boldsymbol{r})^T \Delta \boldsymbol{w} + \frac{\rho}{2} \sum_{i=1}^N (\mathrm{TC}(\Delta w_i))^2 \tag{1}$$

where the first two terms represent the trade-off between the worst-case CVaR and the added expected return caused by rebalancing with the risk-aversion coefficient $\kappa \in \mathbb{R}_+$; and the last one is a sum of the quadratic penalty term of transaction costs with a given penalty parameter $\rho \in \mathbb{R}_+$. The constraint is called ‘ budget constraint’. It means that the added value of portfolio after rebalancing is equivalent to the external capital flow. Notice that, transaction costs are paid out of an external account, and hence externally financed (Guastaroba and Speranza 2012). And that is why we add the penalty term of transaction costs to the objective function of problem (1).

We next reformulate problem (1) via a technical lemma established in Bazaraa *et al.* (2013). For convenience, we denote $p(\boldsymbol{r}) \in \mathcal{P}$ by $\boldsymbol{r} \sim (\mu, \Sigma)$ sometimes.

Lemma 2.1 *Suppose that ξ is a random variable with $\mathbb{E}(\xi) = \mu_\xi$ and $Cov(\xi) = \sigma_\xi^2$. Then, for any real number θ, there holds*

$$\sup_{\xi \sim (\mu_\xi, \sigma_\xi)} \mathbb{E}[(\theta - \xi)_+] = \frac{\theta - \mu_\xi + \sqrt{\sigma_\xi^2 + (\theta - \mu_\xi)^2}}{2}. \tag{2}$$

We assume that the loss function of portfolio rebalancing $f(\Delta \boldsymbol{w}, \boldsymbol{r}) \triangleq \xi = -\boldsymbol{r}^T(\boldsymbol{w}_0 + \Delta \boldsymbol{w})$, then $\xi \sim (\mu_\xi, \sigma_\xi)$ with $\mu_\xi = \mu^T(\boldsymbol{w}_0 + \Delta \boldsymbol{w})$ and $\sigma_\xi^2 = (\boldsymbol{w}_0 + \Delta \boldsymbol{w})^T \Sigma (\boldsymbol{w}_0 + \Delta \boldsymbol{w}))$. It then follows from Lemma 2.1 that we have

$$\begin{aligned}&\sup_{\boldsymbol{r} \sim (\mu, \Sigma)} \mathbb{E}\left[(f(\Delta \boldsymbol{w}, \boldsymbol{r}) - \alpha)_+\right] \\ &\quad = \sup_{\boldsymbol{r} \sim (\mu, \Sigma)} \mathbb{E}\left[\left(-\boldsymbol{r}^T(\boldsymbol{w}_0 + \Delta \boldsymbol{w}) - \alpha\right)_+\right] \\ &\quad = \sup_{\xi \sim (\mu_\xi, \sigma_\xi)} \mathbb{E}\left[(-\xi - \alpha)_+\right] \\ &\quad = \frac{1}{2}\left(\sqrt{\sigma_\xi^2 + (\alpha + \mu_\xi)^2} - (\alpha + \mu_\xi)\right)\end{aligned}$$

Hence, the calculation of the worst-case CVaR can be achieved by solving the following optimization:

$$\min_{\alpha \in \mathbb{R}} \alpha + \frac{1}{2(1-\beta)}\left(\sqrt{\sigma_\xi^2 + (\alpha + \mu_\xi)^2} - (\alpha + \mu_\xi)\right)$$

The optimal solution α^* can be easily obtained as follows:

$$\alpha^* = -\mu_\xi + \frac{2\beta - 1}{2\sqrt{\beta(1-\beta)}} \sigma_\xi,$$

and the corresponding optimal value is the worst-case CVaR, formulated as

$$\mathrm{WCVaR}_\beta(\Delta \boldsymbol{w}) = -\mu_\xi + \sqrt{\frac{\beta}{1-\beta}}\sigma_\xi. \tag{3}$$

Therefore, problem (1) is equivalent to

$$(\mathbf{P}) \quad \min_{\Delta \boldsymbol{w}\in\Omega_1} g(\Delta \boldsymbol{w}) \tag{4}$$

where

$$g(\Delta \boldsymbol{w}) = \frac{\rho}{2}\Delta \boldsymbol{w}^T\tilde{\Sigma}\Delta \boldsymbol{w} + \kappa\tilde{\beta}\sqrt{(\boldsymbol{w}_0+\Delta \boldsymbol{w})^T\Sigma(\boldsymbol{w}_0+\Delta \boldsymbol{w})} - (\kappa+1)\mu^T\Delta \boldsymbol{w}$$

with $\tilde{\Sigma} = \mathcal{D}([\boldsymbol{c}\circ(\mathbf{1}+\boldsymbol{r}_t)]^2)$ and $\tilde{\beta} = \sqrt{\frac{\beta}{1-\beta}}$. We next establish closed-form expressions of problem (**P**) via convex duality theory. We first establish the dual program of problem (**P**), where Theorem 2.2 can be proved via similar argument to that in Appendix 1 of Yang *et al.* (2016).

THEOREM 2.2 *The dual program of problem* (**P**) *(or problem* (1)*) is given as follows:*

$$(\mathbf{P}') \quad \min_{\boldsymbol{y}\in\mathbb{R}^N} \boldsymbol{y}^T\hat{\Sigma}\boldsymbol{y} + \hat{\boldsymbol{c}}^T\boldsymbol{y}$$
$$\text{s.t.} \quad \boldsymbol{y}^T\Sigma^{-1}\boldsymbol{y} \le \frac{1}{4}\kappa\tilde{\beta}. \tag{5}$$

Moreover, the optimal solution $\Delta \boldsymbol{w}^*$ *of problem* (**P**) *can be formulated by the optimal solution* $\boldsymbol{y}^*$ *of problem* (**P**′) *as*

$$\Delta \boldsymbol{w}^* = \frac{1}{\rho}\tilde{\Sigma}^{-1}\left(\lambda^*\mathbf{1} + (\kappa+1)\mu + 2\boldsymbol{y}^*\right) - \boldsymbol{w}_0, \tag{6}$$

where

$$\hat{\Sigma} = \tilde{\Sigma}^{-1} - \frac{\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}};$$
$$\hat{\boldsymbol{c}} = (\kappa+1)\hat{\Sigma}\mu + \rho(S_0+\Delta S)\frac{\tilde{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}};$$
$$\lambda^* = \frac{\rho(S_0+\Delta S) - \mathbf{1}^T\tilde{\Sigma}^{-1}\left((\kappa+1)\mu + 2\boldsymbol{y}^*\right)}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}}.$$

One can observe that problem (**P**) will be solved efficiently if there exists an efficient method for its dual program (**P**′). We next give the closed-form solution for problem (**P**′), whose proof is deferred to Appendix 1.

THEOREM 2.3 *Define the eigenvalue decomposition of* Σ *as*

$$\Sigma = Q\mathcal{D}(\sigma(\Sigma))Q^{-1}.$$

Then, the optimization problem (**P**′) *admits the following closed-form solution*

$$\boldsymbol{y}^*(\eta) = \frac{1}{2}\left[A(\eta) - \frac{A(\eta)\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}A(\eta)}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1} + \mathbf{1}^T\tilde{\Sigma}^{-1}A(\eta)\tilde{\Sigma}^{-1}\mathbf{1}}\right]\hat{\boldsymbol{c}}, \tag{7}$$

where

$$A(\eta) = Q\left[\mathcal{D}\left(1/\boldsymbol{a}(\eta)\right)\right]^\dagger Q^{-1},$$
$$\boldsymbol{a}(\eta) = \eta/\sigma(\Sigma) - 1/(\boldsymbol{c}\circ(\mathbf{1}+\boldsymbol{r}_t))^2,$$

and $\eta > 0$ *is the Lagrangian multiplier corresponding to constraint* (5)*, satisfying*

$$h(\eta) = \boldsymbol{y}^*(\eta)^T\Sigma^{-1}\boldsymbol{y}^*(\eta) - \frac{1}{4}\kappa\tilde{\beta} = 0.$$

We next need to check whether there exists a Lagrangian multiplier $\eta^* > 0$ such that $h(\eta^*) = 0$.

For solving such a Lagrangian multiplier, we utilize the following property of the function $h(\eta)$ (See Appendix 2 for a proof).

THEOREM 2.4 $h(\eta)$ *is continuous and monotonically decreasing with respect to* η*. Moreover, if*

$$\hat{\boldsymbol{c}}^T\hat{\Sigma}^\dagger\Sigma^{-1}\hat{\Sigma}^\dagger\hat{\boldsymbol{c}} > \kappa\tilde{\beta}, \tag{8}$$

then there must exist a Lagrangian multiplier $\eta^* > 0$ *such that* $h(\eta^*) = 0$*. Otherwise, the optimal solution*

$$\boldsymbol{y}^* = -\frac{1}{2}\hat{\Sigma}^\dagger\hat{\boldsymbol{c}}. \tag{9}$$

From Theorem 2.4, the closed-form solution for problem (**P**′) can be obtained via formula (9) if inequality (8) does not hold. Otherwise, we need to search for the root of $h(\eta) = 0$. By fact, it is tractable by means of the monotonic continuity of $h(\eta)$. We can call the secant root-finding (S-RF) algorithm proposed by Dai and Fletcher (2006) to solve equation $h(\eta) = 0$ efficiently. The S-RF algorithm consists of a *bracketing phase* in which an interval $[\underline{\eta}, \overline{\eta}]$ containing a solution is sought, followed by a *secant phase* in which the determined interval is uniformly reduced until the preset tolerance on either the function value or the length of the interval $\Delta\eta$ is met (see Dai and Fletcher 2006, for details). Thereby, followed from (6), we establish a closed-form solution for problem (1).

REMARK 1 We note that an eigenvalue decomposition of Σ is the dominant computation cost for the closed-form solution of problem (1). However, only one eigenvalue decomposition is necessary. For large-scale portfolios (e.g. $N \ge 1000$), we suggest applying the linear-time SVD algorithm developed by Drineas *et al.* (2006) or the sparsing technique of the eigenportfolios addressed in Xiong and Akansu (2019).

3. Industry risk analysis for sparse rebalancing

In this section, we consider a practical financial case,† aided by the analytical solution for problem (**P**), to analyze the industry risk of sparse rebalancing in the absence of diversification constraints.

† Different from the MV-based observation of portfolio diversification in Markowitz 1952, we focus on the cumulative effect of the industry after the robust CVaR rebalancing with cardinality constraint.

Table 1. Industry names and the resulting ticker symbols included.

Industry names	Ticker symbols
Electronic (1)	'002138' (1–1), '002056' (1–2), '600563' (1–3), '000988' (1–4)
Machinery(2)	'300024' (2–1), '002158' (2–2), '600835' (2–3), '600262' (2–4)
Computer (3)	'300033' (3–1), '002253' (3–2), '600536' (3–3), '300042' (3–4)
Food & drink (4)	'600519' (4–1), '002304' (4–2), '000596' (4–3), '600779' (4–4)
Medical biology (5)	'600436' (5–1), '600276' (5–2), '600085' (5–3), '300015' (5–4)

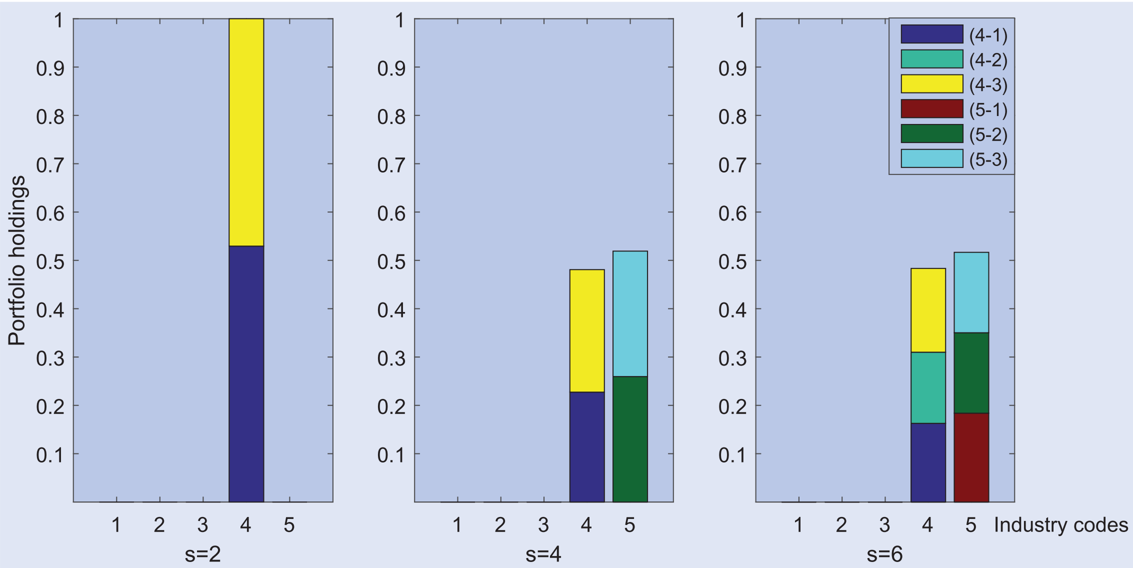

Figure 1. Comparison the global solutions of problem ($\mathbf{P_c}$) with different s_1.

We consider a cardinality-constrained rebalancing strategy within problem ($\mathbf{P}$) by adding to the constraint set Ω_1 a cardinality constraint, written as

$$(\mathbf{P_c}) \quad \min_{\Delta \boldsymbol{w} \in \Omega_1} g(\Delta \boldsymbol{w})$$
$$\text{s.t.} \quad \|\boldsymbol{w}_0 + \Delta \boldsymbol{w}\|_0 \leq s_1. \tag{10}$$

In particular, there have been many effective approximation algorithms for solving problem ($\mathbf{P_c}$), e.g. heuristic algorithm (Beasley *et al.* 2003), Lagrangian relaxation method (Shaw *et al.* 2008), and projected gradient method (Xu *et al.* 2016, 2019), etc. Here, we use the enumeration method to obtain the global optimal solution $\Delta w^*_{P_c}$ of a small-scale problem ($\mathbf{P_c}$) with a given sparsity s_1. That is, we solve $\sum_{1\leq i\leq s_1} C^i_N$ possible cases, where every case can be efficiently solved by means of the closed-form solution of problem ($\mathbf{P}$).

We next study such a financial case. Specially, we select 4 securities from 5 industries, respectively, referring to the first-level industry classification of SWS Index.† The specific industry names and ticker symbols can be seen in table 1, where the numbers in brackets are self-defined in order to make them available easily (e.g.“1” denotes the electronics industry, “1-1” denotes the stock '002138' and so on).

We calculate their log return rates formulated by $\log \frac{p_{j,t}}{p_{j,t-1}}$, where $p_{j,t}$ is the closing price‡ of security j in period t, by using 851 daily closing prices during January 2012 and May 2017.§ Then, we estimate the resulting expectations and covariance matrix of returns with sample moments, respectively. For convenience, we set $S_0 = 1, \Delta S = 0.5, \kappa = 1, \beta = 0.95, \rho = 10, \boldsymbol{c} = 0.001 \cdot \mathbf{1}$ and the initial portfolio is the naive evenly weighted one, i.e. $\boldsymbol{w}_0 = 1/N \cdot \mathbf{1}$. For comparison, we normalize the portfolio holdings after rebalancing, and the output is $(w_0 + \Delta w^*_{P_c})/(S_0 + \Delta S)$. Then, we examine the effects of diversification from the perspective of the industry coverage.

Figure 1 displays the invested securities of optimal portfolio after rebalancing from all five industries. The chart also provides a contrast of the industry coverage cases of problem ($\mathbf{P_c}$) with $s_1 = 2$ against those of $s_1 = 4$ and $s_1 = 6$, respectively. We observe that (i) the optimal portfolio generated by ($\mathbf{P_c}$) is concentrated in the food & drink industry when the investment scale is limited to 2; (ii) the over-concentrated configuration still remains as the upper bound of the investment scale increases since the optimal portfolio after sparse rebalancing is still only concentrated in the food & drink and medical biology industries regardless of the increase of sparsity from 4 to 6. Therefore, by observation, the optimal

† see http://www.swsindex.com/idx0120.aspx?columnid = 8832 for a detailed classification list.

‡ Thereafter, unless explicitly stated, the closing price in the text refers to the 'comparable version' considering cash dividend reinvestment, whose explanation is deferred to Appendix 3 price.

§ Suspension dates included in the collected time series will be removed.

portfolio holdings after sparse rebalancing are overly concentrated in few industries. It implies that their hidden risk should be treated with caution, when there are sudden policy-oriented changes in some investment industry. Next, we will investigate this aspect by adding diversification constraints.

4. Sparse and diversified portfolio rebalancing model

In this section, we attempt to remedy the above hidden industry risk by establishing a sparse and diversified portfolio rebalancing model. Then, we propose a distributed consensus-type ADMM, where each subproblem admits a closed-form solution.

4.1. Rebalancing model with cardinality and diversification constraints

We observe from the above case that there is hidden risk in sparse rebalancing strategies–it is the over-concentration of portfolio holdings in few industries or sectors. We would like to mitigate these shortcoming and maintain the strength of sparse strategies by introducing several financial integer constraints on the basis of problem (1).

We first impose a buy-in range, indicating the minimum (maximum) level below (above) which an asset cannot be traded. That is, a range on the holdings after trading is imposed for reasons like, institutional restrictions (e.g. non-negativity restrictions to reflect no short sales) and/or errors in the estimations of the returns and covariances (Perold 1984). Combined with cardinality constraint (10), we rewrite them as ‘ buy-in threshold constraints’, denoted as

$$\Omega_2 = \left\{ \Delta \boldsymbol{w} \in \mathbb{R}^N : \begin{array}{c} l \cdot \boldsymbol{\delta} \leq \boldsymbol{w}_0 + \Delta \boldsymbol{w} \leq u \cdot \boldsymbol{\delta}; \\ \mathbf{1}^T \boldsymbol{\delta} \leq s_1; \quad \boldsymbol{\delta} \in \{0,1\}^N. \end{array} \right\},$$

where $\delta_j \in \{0, 1\}$ is an extra binary variable, $l = \rho_l(S_0 + \Delta S)$ and $u = \rho_u(S_0 + \Delta S)$ are lower and upper bounds of portfolio holdings, respectively. Generally, investment institution sets bounds as $\rho_l = 0.1\%, \rho_u = 50\%$. But for some special institutions, the bounds may be lower (e.g. $\rho_u = 5\%$ for private equity fund).

Next, we assume there are L economy industries linked with N securities in portfolio, so that there exists an exact partition of $\{1, 2, \ldots, N\}$, written as $\{S_j\}_{j=1,2,\ldots,L}$, affiliated with every industry. Then, L_j is given as the minimum level which ensures that sector j is a diversified industry. A binary variable $\zeta_j \in \{0, 1\}$ is also associated with every economic industry j such that $\zeta_j = 1$ if industry j is diversified; and 0, otherwise. In addition, in order to satisfy the diversification requirement, we introduce a cardinality constraint to ensure that the number of diversified industries is at least s_2. So the ‘ diversification constraints’ can be given as follows:

$$\Omega_3 = \left\{ \Delta \boldsymbol{w} \in \mathbb{R}^N : \begin{array}{c} \mathbf{1}^T [\boldsymbol{w}_0 + \Delta \boldsymbol{w}]_{S_j} \geq L_j \zeta_j, \ j = 1, 2, \ldots, L; \\ \boldsymbol{w}_0 + \Delta \boldsymbol{w} \succeq 0; \ \mathbf{1}^T \boldsymbol{\zeta} \geq s_2; \ \boldsymbol{\zeta} \in \{0,1\}^L. \end{array} \right\}.$$

We now establish a final version for the robust CVaR portfolio rebalancing problem with budget, buy-in threshold and diversification constraints as follows:

$$(\mathbf{P_{cd}}) \quad \min_{\Delta \boldsymbol{w} \in \bigcap_{i=1}^{3} \Omega_i} g(\Delta \boldsymbol{w}).$$

Remark 2 Constraint $\boldsymbol{w}_0 + \Delta \boldsymbol{w} \succeq 0$ in Ω_3 is redundant for problem $(\mathbf{P_{cd}})$. However, it is necessary for saving calculation cost in our designed algorithm (see Subsection 4.2 for details).

4.2. ADMM for problem $(\mathbf{P_{cd}})$

In this section, we propose a distributed consensus-type ADMM for solving problem $(\mathbf{P_{cd}})$, in which each subproblem admits a closed-form solution.

4.2.1. Algorithm framework. We first make the following assumptions throughout the remainder of this paper in order to ensure the feasibility of problem $(\mathbf{P_{cd}})$.

Assumption 1 There exist an integer $m \in [1, s_1]$ and a nonempty set $B \subseteq \{1, 2, \ldots, L\}$ such that

$$\sum_{j \in B} |S_j| \geq m \geq |B| \geq s_2; \tag{11}$$

$$S_0 + \Delta S \in [ml, mu]; \tag{12}$$

$$m \cdot \max_{j \in B} \{L_j\} \leq S_0 + \Delta S. \tag{13}$$

Remark 3 Note that, if (11)–(13) hold, there must exist a feasible solution for problem $(\mathbf{P_{cd}})$ (see Appendix 4 for a detailed explanation). However, not every feasible solution must satisfy the above conditions. For example, assume that there are $N = 12$ securities linked with $L = 6$ economy industries; that is, $S_j = \{2j - 1, 2j\}$ for every $j = 1, \ldots, L$. Set $l = 0.01, u = 0.5, s_1 = 6, S_0 = 1, \ \Delta S = 0.5, \ s_2 = m = 5, \ L_j = 0.4$ for $j = 1, 2, 3, 4$ and $L_j = 0.3$ for $j = 5, 6$. We construct a point $\Delta \boldsymbol{w}$ satisfying $w_{0,i} + \Delta w_i = 0.2$ for $i = 1, 2, \ldots, 8$, $w_{0,9} + \Delta w_9 = 0.3$, and others equal to zero. We can verify easily $\Delta \boldsymbol{w}$ is feasible, but violates (13). It implies the optimal solution of problem $(\mathbf{P_{cd}})$ may not satisfy Assumption 1.

Next, we establish a distributed consensus-type ADMM algorithm framework for solving problem $(\mathbf{P_{cd}})$. Note that this problem can be equivalently reformulated as

$$\min_{\Delta \boldsymbol{w}_i \in \Omega_i, \boldsymbol{z} \in \mathbb{R}^N} \{ g(\Delta \boldsymbol{w}_1) : \Delta \boldsymbol{w}_i = \boldsymbol{z}, \ i = 1, 2, 3 \}. \tag{14}$$

Its augmented Lagrangian function is given by

$$\begin{aligned} &\mathcal{L}_\theta(\Delta \boldsymbol{w}_1, \Delta \boldsymbol{w}_2, \Delta \boldsymbol{w}_3, \boldsymbol{z}; \boldsymbol{y}) \\ &\quad = g(\Delta \boldsymbol{w}_1) + \sum_{i=1}^{3} \left(\boldsymbol{y}_i^T (\Delta \boldsymbol{w}_i - \boldsymbol{z}) + \frac{\theta}{2} \|\Delta \boldsymbol{w}_i - \boldsymbol{z}\|_2^2 \right), \end{aligned} \tag{15}$$

where $\boldsymbol{y} = (\boldsymbol{y}_1; \boldsymbol{y}_2; \boldsymbol{y}_3) \in \mathbb{R}^{3N}$, $\theta > 0$ is a penalty parameter. The resulting iterative scheme of ADMM follows:

$$\begin{cases} \Delta \boldsymbol{w}_i^{k+1} & := \arg\min_{\Delta \boldsymbol{w}_i \in \Omega_i} \mathcal{L}_\theta(\Delta \boldsymbol{w}_i, \boldsymbol{z}^k; \boldsymbol{y}^k); \\ \boldsymbol{z}^{k+1} & := \Delta \bar{\boldsymbol{w}}^{k+1} + \bar{\boldsymbol{y}}^k / \theta; \\ \boldsymbol{y}_i^{k+1} & := \boldsymbol{y}_i^k + \theta(\Delta \boldsymbol{w}_i^{k+1} - \boldsymbol{z}^{k+1}), \end{cases} \tag{16}$$

where $\Delta\bar{\boldsymbol{w}}^k = \frac{1}{3}\sum_{i=1}^{3}\Delta\boldsymbol{w}_i^k, \bar{\boldsymbol{y}}^k = \frac{1}{3}\sum_{i=1}^{3}\boldsymbol{y}_i^k$. One can observe that the first and last steps are carried out independently for each $i = 1, 2, 3$. The processing variable z that handles the global solution is sometimes called the *central collector* or the *fusion center*. Then, we take a popular termination criterion of consensus ADMM proposed by Boyd *et al.* (2011), which is the primal residual $\boldsymbol{r}^k = (\boldsymbol{r}_1^k; \boldsymbol{r}_2^k; \boldsymbol{r}_3^k)$ and the dual residual $\boldsymbol{s}^k = (\boldsymbol{s}_1^k; \boldsymbol{s}_2^k; \boldsymbol{s}_3^k)$ must be small, satisfying

$$\|\boldsymbol{r}^k\|_2^2 = \sum_{i=1}^{3}\|\Delta\boldsymbol{w}_i^k - \Delta\bar{\boldsymbol{w}}^k\|_2^2 \le \epsilon^{pri};$$
$$\|\boldsymbol{s}^k\|_2^2 = 3\theta^2\|\Delta\bar{\boldsymbol{w}}^k - \Delta\bar{\boldsymbol{w}}^{k-1}\|_2^2 \le \epsilon^{dual},$$

with

$$\boldsymbol{r}_i^k = \Delta\boldsymbol{w}_i^k - \Delta\bar{\boldsymbol{w}}^k,\ \ \boldsymbol{s}_i^k = -\theta(\Delta\bar{\boldsymbol{w}}^k - \Delta\bar{\boldsymbol{w}}^{k-1}),\ i = 1, 2, 3;$$
$$\epsilon^{pri} = \sqrt{N}\epsilon^{abs} + \epsilon^{rel}\max_i\{\|\Delta\boldsymbol{w}_i^k\|_2\};$$
$$\epsilon^{dual} = \sqrt{N}\epsilon^{abs} + \epsilon^{rel}\sum_{i=1}^{3}\|\boldsymbol{y}_i^k\|_2,$$

where $\epsilon^{abs} > 0$ is an absolute tolerance and $\epsilon^{rel} > 0$ is a relative tolerance. In addition, some recent literatures showed varying penalty parameter θ might accelerate the convergence rate of ADMM (He *et al.* 2000, Wang and Liao 2001), although there still exists no theoretical guarantee so far. We adopt a simple scheme for θ that often works well:

$$\theta^{k+1} := \begin{cases} \nu^{incr}\theta^k & \text{if } \|\boldsymbol{r}^k\|_2^2 > \gamma\|\boldsymbol{s}^k\|_2^2; \\ \theta^k/\nu^{decr} & \text{if } \|\boldsymbol{s}^k\|_2^2 > \gamma\|\boldsymbol{r}^k\|_2^2; \\ \theta^k & \text{otherwise,} \end{cases}$$

where $\gamma > 1$, $\nu^{incr} > 1$, and $\nu^{decr} > 1$ are parameters.

4.2.2. The closed-form solutions for subproblems. We next establish the closed-for solution of $\Delta\boldsymbol{w}_i (i = 1, 2, 3)$, respectively. The $\Delta\boldsymbol{w}_1^{k+1}$-iteration is to solve:

$$\min_{\Delta\boldsymbol{w}_1\in\Omega_1} g'(\Delta\boldsymbol{w}_1; \boldsymbol{y}_1^k, \boldsymbol{z}^k), \tag{17}$$

where

$$g'(\Delta\boldsymbol{w}_1; \boldsymbol{y}_1^k, \boldsymbol{z}^k) = \frac{\rho}{2}\Delta\boldsymbol{w}_1^T\tilde{\Sigma}'\Delta\boldsymbol{w}_1 + \kappa\tilde{\beta}\sqrt{(\boldsymbol{w}_0 + \Delta\boldsymbol{w}_1)^T\Sigma(\boldsymbol{w}_0 + \Delta\boldsymbol{w}_1)} - (\kappa + 1)(\mu')^T\Delta\boldsymbol{w}_1,$$

with

$$\tilde{\Sigma}' = \tilde{\Sigma} + \frac{\theta^k}{\rho}\boldsymbol{I};\ \ \mu' = \mu + \frac{1}{\kappa + 1}\left(\theta^k\boldsymbol{z}^k - \boldsymbol{y}_1^k\right).$$

It is not hard to see problem (17) is similar to problem (**P**), of which the closed-form solution can be achieved via formulas (6)–(9).

Introduce the following notations.

- $\boldsymbol{a}^k = \boldsymbol{z}^k - \frac{1}{\theta^k}\boldsymbol{y}_2^k - \boldsymbol{w}_0; \boldsymbol{b}^k = \boldsymbol{z}^k - \frac{1}{\theta^k}\boldsymbol{y}_3^k - \boldsymbol{w}_0; \boldsymbol{d} \in \mathbb{R}^L$ with $d_j = \mathbf{1}^T[\boldsymbol{b}^k]_{S_j}$;
- $J_1 = \{j : d_j \ge L_j\};\ J_2 = \{j : 0 < d_j < L_j\};\ J_3 = \{j : d_j \le 0\}$.

Then, the $\Delta\boldsymbol{w}_2^{k+1}$-iteration can be written

$$\min_{\Delta\boldsymbol{w}_2\in\Omega_2} \|\Delta\boldsymbol{w}_2 + \boldsymbol{w}_0 - \boldsymbol{a}^k\|_2^2 \tag{18}$$

The resulting closed-from solution can be given (see Appendix 5 for a proof):

THEOREM 4.1 *Problem* (18) *has the following closed-form solution*

$$[\Delta\boldsymbol{w}_2^{k+1}]_j = \begin{cases} [\Pi_{[l,u]}(\boldsymbol{a}^k)]_j - w_{0,j}, & \text{if } j \in I_{s_1}(\boldsymbol{a}^k); \\ -w_{0,j}, & \text{otherwise,} \end{cases} \tag{19}$$

Similarly, the $\Delta\boldsymbol{w}_3^{k+1}$-iteration is written

$$\min_{\Delta\boldsymbol{w}_3\in\Omega_3} \|\Delta\boldsymbol{w}_3 + \boldsymbol{w}_0 - \boldsymbol{b}^k\|_2^2 \tag{20}$$

We establish the following closed-form solution (see the proof in Appendix 6):

THEOREM 4.2 *If $s_2 \le |J_1|$, the closed-form solution of problem* (20) *is given*

$$[\Delta\boldsymbol{w}_3^{k+1}]_i = \begin{cases} -w_{0,i}, & \textit{if } i \in S_j \textit{ and } j \in J_3; \\ b_i^k - w_{0,i}, & \textit{otherwise,} \end{cases} \tag{21}$$

If $s_2 > |J_1|$, the closed-form solution is

$$[\Delta\boldsymbol{w}_3^{k+1}]_i = \begin{cases} b_i^k - w_{0,i}, & \textit{if } i \in S_j \textit{ and } j \in J_1 \cup J_2 - \bar{J}; \\ b_i^k - w_{0,i} + cons, & \textit{if } i \in S_j \textit{ and } j \in \bar{J}; \\ -w_{0,i}, & \textit{otherwise,} \end{cases} \tag{22}$$

where $\bar{J} = I_{s_2-|J_1|}([\boldsymbol{d}]_{J_2\cup J_3})$ and $cons = \frac{L_j - d_j}{|S_j|}$ is a constant that depends on index j.

5. Empirical analysis

In this section, we conduct numerical experiments to compare the out-of-sample performance of the sparse and diversified rebalancing strategy with the sparse one and no-rebalancing in order to illustrate the effects of the diversification benefits from the perspective of homo economicus. All the computational tests are performed on a ThinkPad X240s PC (Intel Core i7-4510U, 2.0GHz, 8GB RAM) with Matlab R2015b.

5.1. Data sources

We consider portfolios composed of positions in stock indices of short ($\le$ 1 years), medium (1–2 years) and long (2–5 years)

Table 2. The time periods of the eight data subsets.

Data sets	Type	In-sample			Out-of-sample		
		Begin	End	Days	Begin	End	Days
up-up1	short	18/07/2014	29/12/2014	111	30/12/2014	12/06/2015	110
up-down1	short	09/09/2015	25/11/2015	56	26/11/2015	04/04/2016	55
down-up1	medium	12/06/2015	24/03/2016	191	25/03/2016	30/12/2016	190
down-down1	medium	24/05/2013	20/12/2013	141	23/12/2013	18/07/2014	140
up-up2	medium	10/03/2016	12/01/2017	208	13/01/2017	20/12/2017	208
up-down2	medium	18/07/2014	04/02/2015	136	05/02/2015	24/08/2015	135
down-up2	long	03/11/2010	22/02/2013	561	25/02/2013	12/06/2015	560
down-down2	long	03/11/2010	17/11/2011	256	18/11/2011	04/12/2012	255

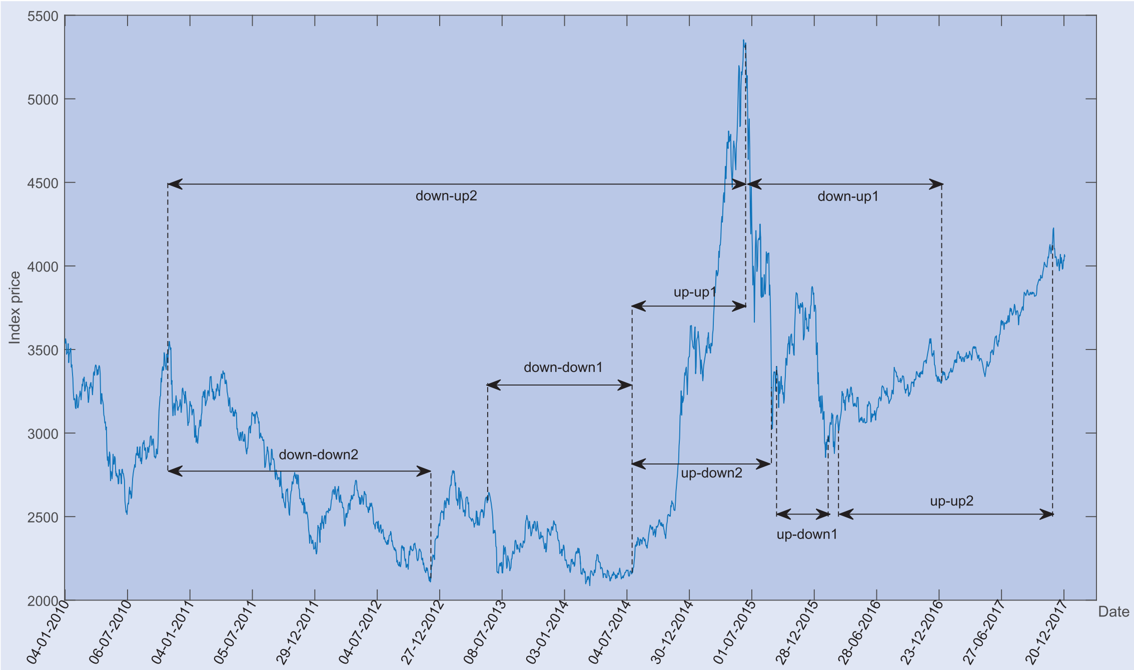

Figure 2. The trends of the market index of CSI300 from 04/01/2010-22/12/2017, and the resulting temporal positioning of eight data sets.

term maturity ranges in the China A-share market, where all the values of the stock indices were obtained from China Stock Market & Accounting Research (CSMAR) database.† And the industry classification criteria, similar to Subsection 3, were referred to the first-level industry classification of SWS Index.

We collect 1939 daily closing prices during January 2010 to December 2017 for 3077 stocks‡ from 28 industries, called as: `all data set`. Referring to the division § in Guastaroba et al. (2009), we take all possible market trends into consideration, and partition `all data set` into 8 data subsets (there are 2 short, 4 medium, and 2 long ones) corresponding to different in-sample and out-of-sample time periods. That is, we equally partition each data subset into an in-sample set and an out-of-sample set, and characterize a market trend¶ going up or down in the in-sample and out-of-sample periods. This is written as: `up-up`, `up-down`, `down-up`, `down-down`. The time period covered by each data subset is summarized in table 2, whereas the resulting temporal positioning is shown in figure 2.

5.2. *Experiment design*

The sparse and diversified rebalancing strategy is obtained by solving problem ($\mathbf{P_{cd}}$), while the sparse rebalancing one is modeled as

$$(\mathbf{P}'_{\mathbf{c}}) \quad \min_{\Delta \boldsymbol{w} \in \bigcap_{i=1}^{2} \Omega_i} \quad g(\Delta \boldsymbol{w}).$$

No-rebalancing means keeping the initial portfolio (the naive evenly weighted portfolio), which belongs to the

† see http://www.gtarsc.com/Home for a detailed introduction.
‡ The total number of the China A-share market is 3489, where 412 special treatment (ST) stocks were removed. This is because ST stocks refer to stocks whose domestic listed companies have suffered losses for two consecutive years and have been warned of delisting risks.
§ Considering the different (up or down) trends in the in-sample and the out-of-sample periods, Guastaroba *et al.* (2009) characterize 4 types of data sets: `up-up`, `up-down`, `down-up` and `down-down`. For example, if the trends go up in both the in-sample and the out-of-sample periods, we call the resulting data set as `up-up`, similar to `up-down`, `down-up` and `down-down`.

¶ Here, we use the closing price trends of CSI 300 index to represent the market trends.

Table 3. The out-of-sample results of WCVaRs, annualized returns, transaction costs on the eight data subsets.

Datasets	ADMM			HEA			No-rebalancing	
	WCVaR	AR	TC	WCVaR	AR	TC	WCVaR	AR
up-up1	9.900e − 3	2.9381	1.962e − 3	7.090e − 3	2.3357	1.956e − 2	7.993e − 3	2.3646
up-up2	1.132e − 3	1.0197	1.974e − 3	1.617e − 3	0.5683	1.944e − 2	1.132e − 3	0.7764
up-down1	1.202e − 2	− 3.9717	2.014e − 3	1.366e − 2	− 7.9374	2.029e − 2	7.863e − 3	− 4.0297
up-down2	2.028e − 3	1.7188	1.991e − 3	2.034e − 3	0.5569	1.981e − 3	8.225e − 3	0.9474
down-up1	1.145e − 3	2.4357	1.972e − 3	1.545e − 3	0.5786	1.956e − 3	2.790e − 3	0.9031
down-up2	1.021e − 3	0.1763	1.992e − 3	1.070e − 3	0.1326	1.993e − 3	2.157e − 3	0.1803
down-down1	6.591e − 4	0.1347	1.989e − 3	5.261e − 4	0.6348	1.958e − 3	1.130e − 3	1.8794
down-down2	8.959e − 4	0.1306	1.991e − 3	1.894e − 3	− 0.1794	1.940e − 3	6.171e − 4	0.0715

buy-and-hold strategy. Then, the experiment will be processed in two stages below.

`Stage 1 (In-sample calculations)` Via the expectation and covariance matrix of in-sample returns, we compute optimal portfolios after rebalancing of problems $(\mathbf{P}_\mathbf{c}')$ and $(\mathbf{P}_\mathbf{cd})$, respectively. For problem $(\mathbf{P}_\mathbf{c}')$, we adopt the hybrid evolutionary algorithm (HEA) framework in Ruiz-Torrubiano and Suárez (2009), where the fitness value (subproblem) is solved by using the Matlab software package CVX (http://cvxr.com/cvx/). For problem $(\mathbf{P}_\mathbf{cd})$, we obtain the optimal solution via the ADMM proposed in Section 4.2. In the tests, apart from the same initial parameter settings as those in Subsection 3, we also set $\rho_l = 1\%, \rho_u = 50\%$, and $s_1 = 20$. For the HEA, we set the initial population size to 100, mutation probability to 1%, cross probability to 50%, and maximum iterations to 10, 000. For the ADMM, we set $s_2 = 5, L_j = 0.1 * (S_0 + \Delta S)$ for every j,† $\theta_0 = 10^2, \epsilon^{abs} = \epsilon^{rel} = 10^{-6}, \gamma = 10, \nu_{incr} = \nu_{decr} = 2$, and the initial points of ADMM $z = 0.01 * \mathbf{ones}(N, 1)$ and $y_i = \mathbf{rand}(N, 1)$ for each i.

`Stage 2 (Out-of-sample calculations)` The performances of the optimal adjusted portfolios obtained in *Stage 1* are tested on the out-of-sample data subsets. In order to measure the out-of-sample performance, we introduce the following three criteria.

- The superiority index of worst-case CVaR ($\mathrm{SI_{WC}}$):

$$\mathrm{SI_{WC}}(A, B) = \frac{\mathrm{WCVaR}_\beta(A) - \mathrm{WCVaR}_\beta(B)}{abs(\mathrm{WCVaR}_\beta(B))} \times 100\%,$$

  where $abs(\cdot)$ is the absolute function, $\mathrm{WCVaR}_\beta(A)$ and $\mathrm{WCVaR}_\beta(B)$ are the out-of-sample worst-case CVaR values solved by methods A and B, respectively.
- The superiority index of annualized return ($\mathrm{SI_{AR}}$):

$$\mathrm{SI_{AR}}(A, B) = \frac{\mathrm{Nr}_A - \mathrm{Nr}_B}{abs(\mathrm{Nr}_B)} \times 100\%,$$

  where $\mathrm{Nr}_A = (\boldsymbol{w}_0 + \Delta \boldsymbol{w}_A^*)^T \mu_{out} \times 25000\%$ is the out-of-sample annualized return of a portfolio obtained by method A, $\Delta \boldsymbol{w}_A^*$ is the optimal traded portfolio obtained by method A, and μ_{out} is the vector of out-of-sample expected returns.
- The industry coverage:

$$\mathrm{Ic}(A) = \frac{n_d}{\mathrm{n}} \times 100\%,$$

  where n_d is the out-of-sample number of diversified industries of the portfolio obtained by method A, and n is the total number of the industries.

REMARK 4 From the definition of $\mathrm{SI_{WC}}$, we can see that if $\mathrm{SI_{WC}} < 0$, $\mathrm{WCVaR}_\beta(A)$ is less than $\mathrm{WCVaR}_\beta(B)$, which implies that the portfolio rebalanced by method A is superior to that by method B in terms of the worst-case CVaR. Otherwise, method A is superior to method B. Similarly, the opposite is true for $\mathrm{SI_{AR}}$.

5.3. *Out-of-sample performance*

For the aforementioned eight different types of data subsets, we compare the out-of-sample performances of $(\mathbf{P}_\mathbf{cd})$ with that of $(\mathbf{P}_\mathbf{c}')$ and no-rebalancing, respectively. Numerical results are presented in tables 3 and 4 and figure 3, by using the normalized portfolio holdings after rebalancing, similar to those in Subsection 3. In particular, we report in table 3 the worst-case CVaR (WCVaR), the annualized return (AR), and transaction costs (TC) by the three strategies. Based on these results, we calculate our proposed three criteria measuring the out-of-sample performance, presented in table 4, along with the comparisons of the corresponding industry holdings after rebalancing, which are reported in figure 3. The abscissa represents the self-defined industry numbers,‡ and the ordinate represents total portfolio holdings of an industry.

From tables 3 and 4, we make the following observations.

(i) Our proposed ADMM strategy ($(\mathbf{P}_\mathbf{cd})$ with ADMM) is superior to $(\mathbf{P}_\mathbf{c}')$ with HEA and no-rebalancing strategies in terms of the values of worst-case CVaR and annualized return. This is because $\mathrm{SI_{WC}}(A, B) < 0$ holds for 75.0% (6/8) instances; $\mathrm{SI_{WC}}(A, C) < 0$ holds for 62.5% (5/8) instances; $\mathrm{SI_{AR}}(A, B) > 0$ holds for 100.0% (8/8) instances; and $\mathrm{SI_{AR}}(A, C) > 0$ holds for 75.0% (6/8) instances.

† It implies that a industry is called as a diversified one if the corresponding industry holding accounts for more than 10%.

‡ A list of industries to which the numbers refer can be seen Appendix 7.

Table 4. The out-of-sample comparison of $\mathrm{SI_{WC}}$, $\mathrm{SI_{AR}}$ and Ic on the eight data subsets.

Datasets	$(\mathbf{P_{cd}})$ VS. $(\mathbf{P'_c})$		$(\mathbf{P_{cd}})$ VS. No-rebalancing		Industry coverage	
	$\mathrm{SI_{WC}}(A,B)$[a]	$\mathrm{SI_{AR}}(A,B)$	$\mathrm{SI_{WC}}(A,C)$[b]	$\mathrm{SI_{AR}}(A,C)$	Ic(A)	Ic(B)
up-up1	0.3962	0.2579	0.2386	0.2455	25.00%	14.29%
up-up2	− 0.3000	0.7943	− 0.0001	0.4282	17.86%	14.29%
up-down1	− 0.1203	0.4996	0.5283	0.0144	17.86%	10.71%
up-down2	− 0.0028	2.0865	− 0.7534	0.8142	17.86%	14.29%
down-up1	− 0.2590	3.2100	− 0.5896	1.6971	17.86%	7.14%
down-up2	− 0.0463	0.3289	− 0.5269	− 0.0226	17.86%	10.71%
down-down1	0.2529	0.7876	− 0.4168	− 0.3962	10.71%	14.29%
down-down2	− 0.5271	1.7284	0.4518	0.8271	14.29%	10.71%

[a]'A' denotes problem $(\mathbf{P_{cd}})$ with our proposed ADMM; 'B' denotes problem $(\mathbf{P'_c})$ with the HEA.[b]'C' denotes no-rebalancing with the naive evenly weighted portfolio.

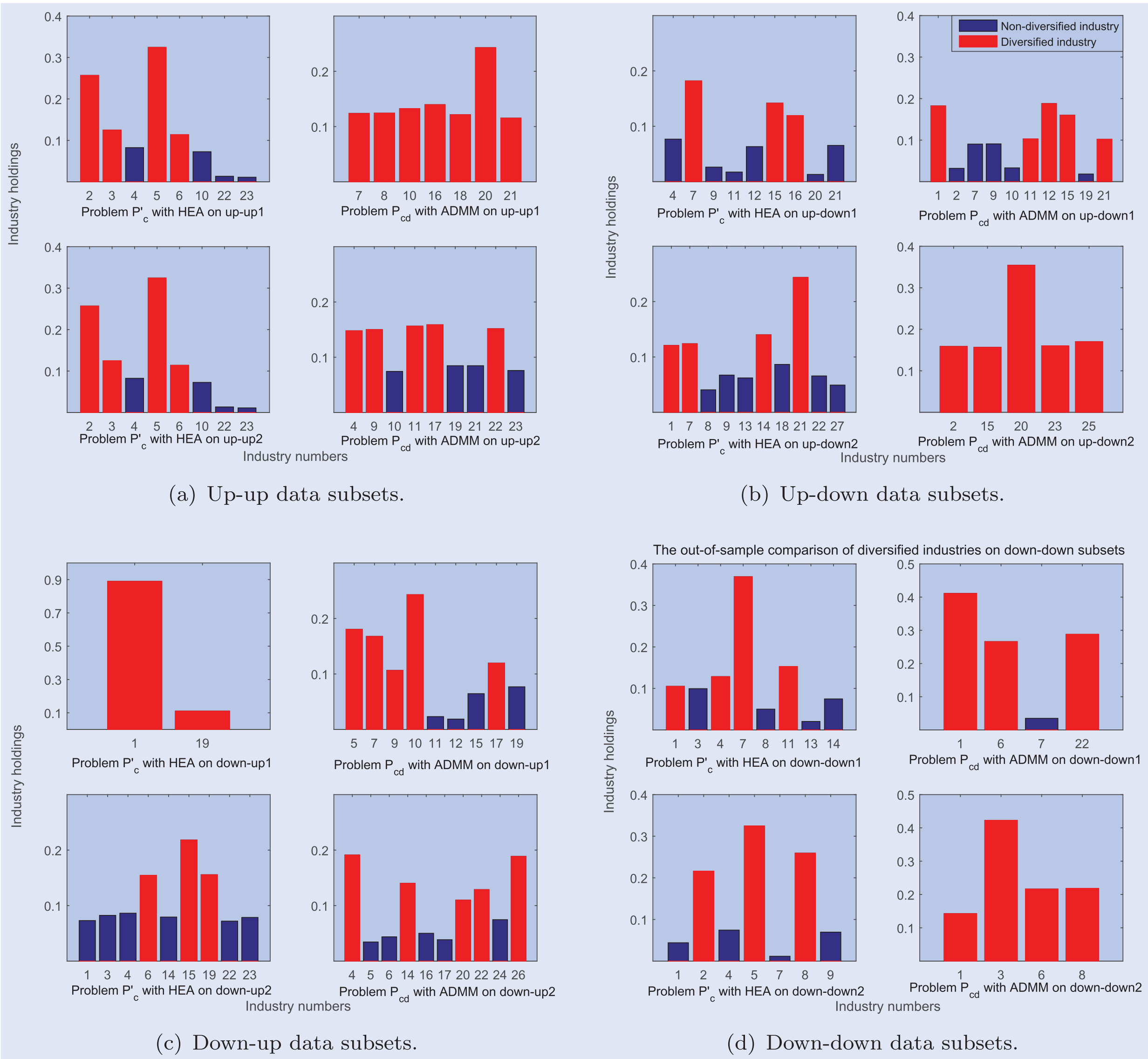

Figure 3. The out-of-sample comparison of diversified industries on the eight data subsets. (a) Up-up data subsets. (b) Up-down data subsets. (c) Down-up data subsets and (d) Down-down data subsets.

(ii) Moreover, since the values of TC are indistinguishable, our ADMM strategy has near identical transactions costs as $(\mathbf{P'_c})$ with HEA.

(iii) However, from the comparison of $\mathrm{SI_{WC}}(A,\cdot)$, $\mathrm{SI_{AR}}(A,\cdot)$, and $Ic(\cdot)$ on the two down-down data subsets, we found that no strategy has obvious advantages for a continuously downward. It implies that this may be a bad moment to enter the market. Temporarily staying out of the market seems to be more sensible.

Combined with table 4 and figure 3, we can make the following observations.

(iv) The portfolio after rebalancing by our ADMM strategy has more diversified industries than that by ($\mathbf{P}_\mathbf{c}'$) with HEA since $\text{Ic}(A) > \text{Ic}(B)$ holds for 87.5% (7/8) instances.
(v) The adjusted portfolio by ($\mathbf{P}_\mathbf{c}'$) with HEA may be overly concentrated in one industry, seen from the industry distribution result of down-up1 subset—the subpicture on the upper left of figure 3(c). On the contrary, our ADMM strategy is able to balance the diversification across investment industries and the sparseness of the portfolio very well in many cases, specifically for up-up1, (seen from the subpicture on the upper right of figure 3(a)), up-down2 (the subpicture on the lower right of figure 3(b)) and down-down2 (the subpicture on the lower right of figure 3(d)) subsets.
(vi) The outcomes of down-down data sunsets generated by our ADMM strategy are not feasible in terms of the constraint set of problem ($\mathbf{P}_\mathbf{cd}$) since the numbers of diversified industries are both less than $s_2 = 5$. Even so, the out-of-sample performance of our strategy is still no worse than that of the other two strategies.

6. Concluding remarks

In this paper, a new portfolio rebalancing problem was proposed with both embedded sparsity and diversification, two desirable properties in practice. The portfolio rebalancing problem was formulated as a mixed-integer SOCP with 'double' cardinality constraints (one is the diversification constraint, and the other is the buy-in threshold constraint). An effective ADMM was proposed to solve the model, where each block has a closed-form solution.

In addition, we illustrated the disadvantages of sparse rebalancing through a practical financial case in the absence of diversification constraints. Then, using actual market data, we presented the computational results of our proposed strategy and a comparison of these results to the standard sparse rebalancing and no-rebalancing strategies. The results demonstrate the superior out-of-sample performance of our proposed model and algorithm. Our sparse and diversified strategy is able to trade off between the limits of investment scale and diversified industry coverage very well, making the portfolio after rebalancing easy to manage and risk resistant.

Acknowledgments

The authors would like to thank the two anonymous referees for their constructive comments which substantially improved the presentation of the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This author is supported by National Natural Science Foundations of China [grant number 11571271], [grant number 11101325] and Key Program of National Natural Science Foundation of China [grant number 11631013]. This author is supported by supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number 06446], and National Natural Science Foundations (NNSF) of China [grant number 11771386], [grant number 11728104]. This author is supported by National Natural Science Foundations of China [grant number 71501155], [grant number 11601409].

References

Artzner, P., Delbaen, F., Eber, J.M. and Heath, D., Coherent measures of risk. *Math. Finance*, 1999, **9**, 203–228.
Basak, S. and Shapiro, A., Value-at-risk-based risk management: Optimal policies and asset prices. *Rev. Financ. Stud.*, 2001, **14**, 371–405.
Bazaraa, M.S., Sherali, H.D. and Shetty, C.M., *Nonlinear Programming: Theory and Algorithms*, 2013 (John Wiley & Sons: Hoboken, NJ).
Beasley, J.E., Meade, N. and Chang, T.J., An evolutionary heuristic for the index tracking problem. *Eur. J. Oper. Res.*, 2003, **148**, 621–643.
Blumensath, T. and Davies, M.E., Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.*, 2009, **27**, 265–274.
Bonami, P. and Lejeune, M.A., An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Oper. Res.*, 2009, **57**, 650–670.
Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J., Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 2011, **3**, 1–122.
Branda, M., Diversification-consistent data envelopment analysis with general deviation measures. *Eur. J. Oper. Res.*, 2013, **226**, 626–635.
Branda, M., Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega*, 2015, **52**, 65–76.
Brodie, J., Daubechies, I., De Mol, C., Giannone, D. and Loris, I., Sparse and stable Markowitz portfolios. *Proc. Nat. Acad. Sci.*, 2009, **106**, 12267–12272.
Bruder, B., Gaussel, N., Richard, J.C. and Roncalli, T., Regularization of portfolio allocation. Available at SSRN 2767358, 2013.
Chang, T.J., Meade, N., Beasley, J.E. and Sharaiha, Y.M., Heuristics for cardinality constrained portfolio optimisation. *Comput. Oper. Res.*, 2000, **27**, 1271–1302.
Chen, X., Xu, F. and Ye, Y., Lower bound theory of nonzero entries in solutions of $l_2 - l_p$ minimization. *SIAM. J. Sci. Comput.*, 2010, **32**, 2832–2852.
Chen, L., He, S. and Zhang, S., Tight bounds for some risk measures, with applications to robust portfolio selection. *Oper. Res.*, 2011, **59**, 847–865.
Chen, C., Li, X., Tolman, C., Wang, S. and Ye, Y., Sparse portfolio selection via quasi-norm regularization. arXiv preprint arXiv:1312.6350, 2013.
Chen, X., Ge, D., Wang, Z. and Ye, Y., Complexity of unconstrained $l_2 - l_p$ minimization. *Math. Program.*, 2014, **143**, 371–383.
Chiodi, L., Mansini, R. and Speranza, M.G., Semi-absolute deviation rule for mutual funds portfolio selection. *Ann. Oper. Res.*, 2003, **124**, 245–265.
Dai, Y.H. and Fletcher, R., New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. *Math. Program.*, 2006, **106**, 403–421.

DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *Rev. Financ. Stud.*, 2007, **22**, 1915–1953.

DeMiguel, V., Garlappi, L., Nogales, F.J. and Uppal, R., A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Manage. Sci.*, 2009, **55**, 798–812.

Drineas, P., Kannan, R. and Mahoney, M.W., Fast Monte Carlo algorithms for matrices II: Computing a low-rank approximation to a matrix. *SIAM J. Comput.*, 2006, **36**, 158–183.

Fabozzi, F.J., Huang, D. and Zhou, G., Robust portfolios: Contributions from operations research and finance. *Ann. Oper. Res.*, 2010, **176**, 191–220.

Fan, J., Zhang, J. and Yu, K., Asset allocation and risk assessment with gross exposure constraints for vast portfolios. arXiv preprint arXiv:0812.2604, 2008.

Fang, Y., Lai, K. and Wang, S., Portfolio rebalancing with transaction costs and a minimal purchase unit. *Dyn. Contin. Discr. Impulsive Syst.*, 2005, **4**, 499–515.

Gaivoronski, A.A. and Pflug, G., Value-at-risk in portfolio optimization: Properties and computational approach. *J. Risk*, 2005, **7**, 1–31.

Gao, J. and Li, D., Optimal cardinality constrained portfolio selection. *Oper. Res.*, 2013, **61**, 745–761.

Ghaoui, L.E., Oks, M. and Oustry, F., Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Oper. Res.*, 2003, **51**, 543–556.

Guastaroba, G. and Speranza, M.G., Kernel search: An application to the index tracking problem. *Eur. J. Oper. Res.*, 2012, **217**, 54–68.

Guastaroba, G., Mansini, R. and Speranza, M.G., Models and simulations for portfolio rebalancing. *Comput. Econom.*, 2009, **33**, 237.

He, Z., Sensitivity estimation of conditional value at risk using randomized quasi-Monte Carlo. arXiv preprint arXiv:1908.07232, 2019.

He, B., Yang, H. and Wang, S., Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *J. Optim. Theory. Appl.*, 2000, **106**, 337–356.

Hong, L.J. and Liu, G., Simulating sensitivities of conditional value at risk. *Manage. Sci.*, 2009, **55**, 281–293.

Konno, H. and Yamazaki, H., Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Manage. Sci.*, 1991, **37**, 519–531.

Kyrillidis, A., Becker, S., Cevher, V. and Koch, C., Sparse projections onto the simplex. In Proceedings of the International Conference on Machine Learning, pp. 235–243, 2013.

Ledoit, O. and Wolf, M., Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance*, 2003, **10**, 603–621.

Ledoit, O. and Wolf, M., A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.*, 2004, **88**, 365–411.

Lejeune, M.A., A VaR Black–Litterman model for the construction of absolute return fund-of-funds. *Quant. Finance*, 2011, **11**, 1489–1501.

Lobo, M.S. and Boyd, S., The worst-case risk of a portfolio. Unpublished manuscript, 2000.

Markowitz, H., Portfolio selection. *J. Finance*, 1952, **7**, 77–91.

Markowitz, H.M., *Portfolio Selection: Efficient Diversification of Investments*, 1970 (Yale University Press: London).

Natarajan, K. and Linyi, Z., A mean–variance bound for a three-piece linear function. *Probab. Engin. Inform. Sci.*, 2007, **21**, 611–621.

Natarajan, K., Sim, M. and Uichanco, J., Tractable robust expected utility and risk models for portfolio optimization. *Math. Finance*, 2010, **20**, 695–731.

Perold, A.F., Large-scale portfolio optimization. *Manage. Sci.*, 1984, **30**, 1143–1160.

Rockafellar, R.T. and Uryasev, S., Optimization of conditional value-at-risk. *J. Risk*, 2000, **2**, 21–42.

Rockafellar, R.T. and Uryasev, S., Conditional value-at-risk for general loss distributions. *J. Bank. Financ.*, 2002, **26**, 1443–1471.

Rockafellar, R.T., Uryasev, S. and Zabarankin, M., Optimality conditions in portfolio analysis with general deviation measures. *Math. Program.*, 2006, **108**, 515–540.

Ruiz-Torrubiano, R. and Suárez, A., A hybrid optimization approach to index tracking. *Ann. Oper. Res.*, 2009, **166**, 57–71.

SantAnna, L.R., Filomena, T.P., Guedes, P.C. and Borenstein, D., Index tracking with controlled number of assets using a hybrid heuristic combining genetic algorithm and non-linear programming. *Ann. Oper. Res.*, 2017, **258**, 849–867.

Schwarz, G., Estimating the dimension of a model. *Ann. Stat.*, 1978, **6**, 461–464.

Sharpe, W.F., Mean-absolute-deviation characteristic lines for securities and portfolios. *Manage. Sci.*, 1971, **18**, B–1.

Shaw, D.X., Liu, S. and Kopman, L., Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. *Optim. Methods Softw.*, 2008, **23**, 411–420.

Sherman, J. and Morrison, W.J., Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Ann. Math. Stat.*, 1950, **21**, 124–127.

Speranza, M.G., Linear programming models for portfolio optimization, 1993.

Strub, O. and Baumann, P., Optimal construction and rebalancing of index-tracking portfolios. *Eur. J. Oper. Res.*, 2018, **264**, 370–387.

Takeda, A., Niranjan, M., Gotoh, J.y. and Kawahara, Y., Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios. *Comput. Manage. Sci.*, 2013, **10**, 21–49.

Tibshirani, R., Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodological)*, 1996, **58**, 267–288.

Wang, S. and Liao, L., Decomposition method with a variable parameter for a class of monotone variational inequality problems. *J. Optim. Theory. Appl.*, 2001, **109**, 415–429.

Woodside-Oriakhi, M., Lucas, C. and Beasley, J.E., Heuristic algorithms for the cardinality constrained efficient frontier. *Eur. J. Oper. Res.*, 2011, **213**, 538–550.

Woodside-Oriakhi, M., Lucas, C. and Beasley, J.E., Portfolio rebalancing with an investment horizon and transaction costs. *Omega*, 2013, **41**, 406–420.

Xiong, A. and Akansu, A.N., On sparsity of eigenportfolios to reduce transaction cost. *J. Capital Markets Stud.*, 2019, **3**, 82–90.

Xu, F., Xu, Z. and Xue, H., Sparse index tracking based on $L\_\{1/2\}$ model and algorithm. arXiv preprint arXiv:1506.05867, 2015.

Xu, F., Lu, Z. and Xu, Z., An efficient optimization approach for a cardinality-constrained index tracking problem. *Optim. Method. Softw.*, 2016, **31**, 258–271.

Xu, F., Dai, Y., Zhao, Z. and Xu, Z., Efficient projected gradient methods for cardinality constrained optimization. *Sci. China Math.*, 2019, 1–24.

Yang, Y., Ahipasaoglu, S. and Chen, J., On the robustness and sparsity trade-off in mean-variance portfolio selection. Available at SSRN 2873037, 2016.

Yu, J.R. and Lee, W.Y., Portfolio rebalancing model using multiple criteria. *Eur. J. Oper. Res.*, 2011, **209**, 166–175.

Zhu, S. and Fukushima, M., Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.*, 2009, **57**, 1155–1168.

Appendices

Appendix 1. Proof of Theorem 2.3

The augmented Lagrangian function of problem ($\mathbf{P}'$) can be given by

$$\mathcal{L}(\boldsymbol{y};\eta) = \boldsymbol{y}^T\hat{\Sigma}\boldsymbol{y} + \hat{\boldsymbol{c}}^T\boldsymbol{y} - \eta\left(\boldsymbol{y}^T\Sigma^{-1}\boldsymbol{y} - \frac{1}{4}\kappa\tilde{\beta}\right),$$

with the Lagrangian multiplier $\eta > 0$. Then, the partial derivative of each variable can also be obtained, respectively,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{y}} = 2\hat{\Sigma}\boldsymbol{y} + \hat{\boldsymbol{c}} - 2\eta\Sigma^{-1}\boldsymbol{y} = 0; \tag{A1}$$

$$\frac{\partial \mathcal{L}}{\partial \eta} = \boldsymbol{y}^T \Sigma^{-1} \boldsymbol{y} - \frac{1}{4}\kappa\tilde{\beta} = 0. \tag{A2}$$

It follows from equation (A1) that we have

$$\boldsymbol{y}^*(\eta) = \frac{1}{2}\left[\eta\Sigma^{-1} - \hat{\Sigma}\right]^{\dagger} \hat{\boldsymbol{c}}, \tag{A3}$$

Substituting (A3) into (A2), we have

$$h(\eta) \triangleq \boldsymbol{y}^*(\eta)^T \Sigma^{-1} \boldsymbol{y}^*(\eta) - \frac{1}{4}\kappa\tilde{\beta} = 0.$$

Since $\eta\Sigma^{-1} - \hat{\Sigma} = (\eta\Sigma^{-1} - \tilde{\Sigma}^{-1}) + \frac{\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}}$, it is not hard to see that, according to the Sherman–Morrison Theorem (Sherman and Morrison 1950), the rank one correction matrix of $\eta\Sigma^{-1} - \tilde{\Sigma}^{-1}$ is nonsingular if

$$\eta \neq \eta_j \triangleq \frac{\sigma_j(\Sigma)}{(c_j(1 + r_{j,t}))^2}, \quad \forall j = 1, 2, \ldots, N.$$

Without loss of generality, we assume $\eta_i \neq \eta_j \ \forall \ i \neq j$, which are called as breakpoints. The resulting collection is defined as $\mathcal{B} = \{\eta_1, \eta_2, \ldots, \eta_N\}$ in nondecreasing order. Set the eigenvalue decomposition $\Sigma = Q\mathcal{D}(\sigma(\Sigma))Q^{-1}$, then we have

- when $\eta \notin \mathcal{B}$,

$$\begin{aligned}
&\left(\eta\Sigma^{-1} - \hat{\Sigma}\right)^{-1} \\
&= \left(\eta\Sigma^{-1} - \tilde{\Sigma}^{-1} + \frac{\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}}\right)^{-1} \\
&= \left(Q\mathcal{D}\left(\eta/\sigma(\Sigma) - \mathbf{1}/(\boldsymbol{c} \circ (\mathbf{1} + \boldsymbol{r}_t))^2\right) Q^{-1} \right.\\
&\quad \left. + \frac{\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}}\right)^{-1}
\end{aligned}$$

  Setting $\boldsymbol{a}(\eta) = \eta/\sigma(\Sigma) - 1/(\boldsymbol{c} \circ (\mathbf{1} + \boldsymbol{r}_t))^2$, $A(\eta) = Q[\mathcal{D}(1/\boldsymbol{a}(\eta))]^{-1}Q^{-1}$, we have

$$\begin{aligned}
\left(\eta\Sigma^{-1} - \hat{\Sigma}\right)^{-1} &= \left(Q\mathcal{D}\left(1/\boldsymbol{a}(\eta)\right) Q^{-1} + \frac{\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1}}\right)^{-1} \\
&= A(\eta) - \frac{A(\eta)\tilde{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T\tilde{\Sigma}^{-1}A(\eta)}{\mathbf{1}^T\tilde{\Sigma}^{-1}\mathbf{1} + \mathbf{1}^T\tilde{\Sigma}^{-1}A(\eta)\tilde{\Sigma}^{-1}\mathbf{1}}
\end{aligned}$$

- when $\eta \in \mathcal{B}$, we set $A(\eta) = Q[\mathcal{D}(1/\boldsymbol{a}(\eta))]^{\dagger}Q^{-1}$, where

$$[\mathcal{D}\left(1/\boldsymbol{a}(\eta)\right)]^{\dagger}_{i,i} = \begin{cases} 0, & \text{if } a_i(\eta) = 0, \\ a_i(\eta), & \text{otherwise.} \end{cases} \tag{A4}$$

For convenience, we uniformly make $A(\eta) = Q[\mathcal{D}(1/\boldsymbol{a}(\eta))]^{\dagger}Q^{-1}$, and thereby the closed-form solution of problem ($\mathbf{P}'$), formulated as (7), can be obtained. This finishes the proof.

Appendix 2. Proof of Theorem 2.4

First, we prove that $h(\eta)$ is continuous. It is clear that $h(\eta)$ is continuous when $\eta \notin \mathcal{B}$. If $\eta = \eta_j \in \mathcal{B}$ for any given j, we have from (A4)

$$\begin{aligned}
&\lim_{\Delta\eta\to 0}\left[\mathcal{D}\left(1/\boldsymbol{a}(\eta_j + \Delta\eta)\right)\right]^{\dagger}_{i,i} \\
&= \lim_{\Delta\eta\to 0} a_i(\eta_j + \Delta\eta) \\
&= \left.\begin{cases} a_i(\eta_j), & \text{if } i \neq j; \\ 0, & \text{otherwise.} \end{cases}\right\} = \left[\mathcal{D}\left(1/\boldsymbol{a}(\eta_j)\right)\right]^{\dagger}_{i,i}
\end{aligned}$$

Therefore, $[\mathcal{D}(1/\boldsymbol{a}(\eta))]^{\dagger}$ is continuous with respect to $\eta > 0$, which implies that $h(\eta)$ is continuous. Next, we prove the monotonicity of $h(\eta)$. Suppose there exist two Lagrangian multipliers η_1 and η_2, corresponding to the optimal solutions of problem ($\mathbf{P}'$), $\boldsymbol{y}^*(\eta_1)$ and $\boldsymbol{y}^*(\eta_2)$, respectively. Due to symmetry, we only prove the desired result for $\eta_1 \leq \eta_2$. By the optimality of $\boldsymbol{y}^*(\eta_1)$ and $\boldsymbol{y}^*(\eta_2)$, we have

$$\begin{aligned}
\mathcal{L}(\boldsymbol{y}^*(\eta_1); \eta_1) &\leq \mathcal{L}(\boldsymbol{y}^*(\eta_2); \eta_1); \\
\mathcal{L}(\boldsymbol{y}^*(\eta_2); \eta_2) &\leq \mathcal{L}(\boldsymbol{y}^*(\eta_1); \eta_2).
\end{aligned}$$

Together these two inequalities imply that

$$\eta_1 h(\eta_1) + \eta_2 h(\eta_2) \leq \eta_1 h(\eta_2) + \eta_2 h(\eta_1),$$

Or equivalently, $(\eta_1 - \eta_2)(h(\eta_1) - h(\eta_2)) \leq 0$; that is, $h(\eta)$ is decreasing. Finally, we prove the existence of η. It follows from (7) that we have

$$\begin{aligned}
\lim_{\eta\to+\infty} \boldsymbol{y}^*(\eta)^T\Sigma^{-1}\boldsymbol{y}^*(\eta) &= \lim_{\eta\to+\infty} \tilde{\boldsymbol{y}}^*(\eta)^T\Sigma^{-1}\tilde{\boldsymbol{y}}^*(\eta) \\
&= \lim_{\eta\to+\infty} \frac{1}{4\eta^2}\hat{\boldsymbol{c}}^T\Sigma\hat{\boldsymbol{c}} = 0
\end{aligned}$$

where $\tilde{\boldsymbol{y}}^*(\eta) = \frac{1}{2\eta}\Sigma\hat{\boldsymbol{c}}$. It implies that $\lim_{\eta\to+\infty} h(\eta) = -\frac{1}{4}\kappa\tilde{\beta} < 0$.

On the other hand, it is easy to see that $\lim_{\eta\to 0}\boldsymbol{y}^*(\eta) = \hat{\boldsymbol{y}} = -\frac{1}{2}\hat{\Sigma}^{\dagger}\hat{\boldsymbol{c}}$. Then, we have

$$\lim_{\eta\to 0} \boldsymbol{y}^*(\eta)^T\Sigma^{-1}\boldsymbol{y}^*(\eta) = \frac{1}{4}\hat{\boldsymbol{c}}^T\hat{\Sigma}^{\dagger}\Sigma^{-1}\hat{\Sigma}^{\dagger}\hat{\boldsymbol{c}}$$

If inequality (8) holds, it is easy to see that $\lim_{\eta\to 0} h(\eta) > 0$, which implies that there must at least exist a Lagrangian multiplier $\eta^* > 0$ such that $h(\eta^*) = 0$. Otherwise, it means that constraint (5) is redundant for problem ($\mathbf{P}'$), that is, the optimal solution $\boldsymbol{y}^* = -\frac{1}{2}\hat{\Sigma}^{\dagger}\hat{\boldsymbol{c}}$. This completes the proof.

Appendix 3. An explanation for the 'comparable' closing prices

From a time perspective, the closing prices announced by Shenzhen Stock Exchange and Shanghai Stock Exchange are not comparable due to

- the changes of share capital caused by bonus issue, rights issue, and stock split, ect.;
- the adjustment of the Stock Exchanges for the daily closing price.

To address this issue, we use the 'comparable' closing prices, provided by the CSMAR Services, of which the computational basis is the closing price on the first day of listing. Specially, for any security j, the comparable closing price in period t can be formulated as

$$p_{j,t} = p_{j,1} \times \prod_{i=2}^{t}(1 + \bar{r}_{j,i}),$$

where

- $p_{j,1}$ is the closing price of security j on the first day of listing;
- $\bar{r}_{j,t}$ is the arithmetic return of security j in period t considering cash dividend reinvestment.

Then, we can calculate the logarithmic return $r_{j,t} = \log\frac{p_{i,t}}{p_{i,t-1}}$, which implies that $r_{j,t} = \log(1 + \bar{r}_{j,t})$.

Appendix 4. The feasibility of problem ($\mathbf{P_{cd}}$)

We first construct a point $\Delta\boldsymbol{w}$ as follows. Given an exact partition $\{S_1, \ldots, S_L\}$, and s_1, s_2, without loss of generality, we assume that there exists a non-empty set $B \subseteq \{1, 2, \ldots, L\}$ affiliated with $\{S_1 \cdots, S_{|B|}\}$. According to (11), we can keep picking an index i from the set S_i for $i = 1, \ldots, |B|$ without repetition, and set $w_{0,i} + \Delta w_i = (S_0 + \Delta S)/m$, until the number of elements is m ($m \leq s_1$). We then make other elements of $\boldsymbol{w}_0 + \Delta \boldsymbol{w}$ equal to 0. Thus, we obtain a point $\Delta \boldsymbol{w}$. Next, we prove it is feasible. It is obvious that $\mathbf{1}^T(\boldsymbol{w}_0 + \Delta \boldsymbol{w}) = S_0 + \Delta S$, that is, $\Delta \boldsymbol{w} \in \Omega_1$. One can also observe that the cardinality of $\boldsymbol{w}_0 + \Delta \boldsymbol{w}$ is $m \leq s_1$, and the corresponding nonzero elements belong to the range $[l, u]$ because of (12), which implies that $\Delta \boldsymbol{w} \in \Omega_2$. Moreover, it is easy to know from (13) that there holds $L_j \leq \sum_{i \in S_j}(w_{0,i} + \Delta w_i) \leq (S_0 + \Delta S)$ for any $j = 1, \ldots, |B|$. It then follows $|B| \geq s_2$ that we know $\Delta \boldsymbol{w} \in \Omega_3$. Therefore, the point $\Delta \boldsymbol{w}$ we construct is feasible. That is, Assumption 1 ensures the feasibility of problem ($\mathbf{P_{cd}}$).

Appendix 5. Proof of Theorem 4.1

Problem (18) can be solved through the following optimization

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^N} \quad & \|\boldsymbol{x} - \boldsymbol{a}\|_2^2 \\ \text{s.t.} \quad & \|x\|_0 \leq s_1; \\ & l \leq x_i \leq u, \ i \in Supp(\boldsymbol{x}), \end{aligned} \tag{A5}$$

where $Supp(\boldsymbol{x})$ denotes the indices set of nonzero elements of $\boldsymbol{x}$. Let Σ_{s_1} be the set including all s_1 index sets, namely, $\Sigma_{s_1} = \{I : |I| \leq s_1, I \subset \{1, 2, \ldots, n\}\}$. Then, for any index set $I \in \Sigma_{s_1}$ and $x_i \neq 0, \forall i \in I$, model (A5) can be formulated as

$$\begin{aligned} \min \quad & \|(\boldsymbol{x} - \boldsymbol{a})_I\|_2^2 + \|(\boldsymbol{a})_{\bar{I}}\|_2^2 \\ \text{s.t.} \quad & l \leq x_i \leq u, \ i \in I; \\ & x_i = 0, \ i \in \bar{I}, \end{aligned}$$

where $\bar{I}$ is the complementary set of I. The optimal solution is given as follows

$$x_i = \begin{cases} [\Pi_{[l,u]}(\boldsymbol{a})]_i, & \text{if } i \in I; \\ 0, & \text{otherwise.} \end{cases}$$

Thus, the original problem is equivalent to finding an optimal index set in Σ_{s_1}, namely,

$$\begin{aligned} & \operatorname{Arg} \min_{I \in \Sigma_{s_1}} \|(\boldsymbol{x} - \boldsymbol{a})_I\|_2^2 + \|(\boldsymbol{a})_{\bar{I}}\|_2^2 \\ & = \operatorname{Arg} \min_{I \in \Sigma_{s_1}} \|(\Pi_{[l,u]}(\boldsymbol{a}) - \boldsymbol{a})_I\|_2^2 + \|(\boldsymbol{a})_{\bar{I}}\|_2^2 \\ & = \operatorname{Arg} \min_{I \in \Sigma_{s_1}} \|(\Pi_{[l,u]}(\boldsymbol{a}) - \boldsymbol{a})_I\|_2^2 - \|(\boldsymbol{a})_I\|_2^2 + \|\boldsymbol{a}\|_2^2 \\ & = \operatorname{Arg} \max_{I \in \Sigma_{s_1}} \|(\boldsymbol{a})_I\|_2^2 - \|(\Pi_{[l,u]}(\boldsymbol{a}) - \boldsymbol{a})_I\|_2^2 \\ & = \operatorname{Arg} \max_{I \in \Sigma_{s_1}} \sum_{i \in I} (a_i^2 - ([\Pi_{[l,u]}(\boldsymbol{a})]_i - a_i)^2). \end{aligned} \tag{A6}$$

Therefore, model (A5) is equivalent to finding an optimal index set I^* such that

$$I^* \in \operatorname{Arg} \max_{I \in \Sigma_{s_1}} \sum_{i \in I} (a_i^2 - ([\Pi_{[l,u]}(\boldsymbol{a})]_i - a_i)^2).$$

Consider the function $b(t) = t^2 - (\Pi_{[l,u]}(t) - t)^2$ for any $t \in \Re$. Its first derivative is given by

$$b'(t) = 2t - 2(t - \Pi_{[l,u]}(t)) = 2\Pi_{[l,u]}(t) > 0,$$

which implies that $b(t)$ is an increasing function, namely, the optimal index set $I^* = I_{s_1}(\boldsymbol{a})$.

In short, we can admit the closed-form solution of problem (A5) as follows:

$$\boldsymbol{x}_i^* = \begin{cases} [\Pi_{[l,u]}(\boldsymbol{a})]_i, & \text{if } i \in I_{s_1}(\boldsymbol{a}); \\ 0, & \text{otherwise,} \end{cases}$$

Then, we replace $\boldsymbol{x}$ with $\Delta \boldsymbol{w} + \boldsymbol{w}_0$, $\boldsymbol{a}$ with $\boldsymbol{a}^k$, and the formula (19) can be obtained.

Appendix 6. Proof of Theorem 4.2

Set $\boldsymbol{x} = \Delta \boldsymbol{w} + w_0$, and problem (20) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}_+^N, \boldsymbol{y} \in \mathbb{R}^L} \quad & \|\boldsymbol{y} - \boldsymbol{d}\|_2^2 \\ \text{s.t.} \quad & \mathbf{1}^T [\boldsymbol{x}]_{S_j} = y_j, \ y_j \geq L_j \zeta_j, \ j = 1, 2, \ldots, L; \\ & \mathbf{1}^T \boldsymbol{\zeta} \geq s_2; \ \boldsymbol{\zeta} \in \{0, 1\}^L. \end{aligned} \tag{A7}$$

We assume the optimal solution of problem (A7) is $(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

- When $s_2 \leq |J_1|$, the cardinality constraint is naturally satisfied. At this time, problem (A7) is reduced to a simple projection problem. Specially, it is easy to see that $y_j^* = 0$ if $j \in J_3$; otherwise, $y_j^* = d_j$. Thereby, the resulting $x_i^* = 0$ if $i \in S_j$ and $j \in J_3$; otherwise, $x_i^* = b_i^k$. Therefore, formula (21) is achieved.
- When $s_2 > |J_1|$, in addition to the index set J_1, we need to pick $s_2 - |J_1|$ indices from $J_2 \cup J_3$ to project the resulting y_j onto the interval $[L_j, +\infty)$ in order to satisfy the cardinality constraint. And the optimal $s_2 - |J_1|$ indices are the ones of $s_2 - |J_1|$ largest entries of $[\boldsymbol{d}]_{J_2 \cup J_3}$, written as $\bar{J}$. Therefore, we can obtain the optimal $\boldsymbol{y}^*$ by setting $y_j^* = d_j$ for $j \in J_1 \cup J_2 - \bar{J}$, $y_j^* = L_j$ for $j \in \bar{J}$, and $y_j^* = 0$ for $j \in J_3 - \bar{J}$. And the optimal x_i^* is unique for $i \in S_j$ and $j \notin \bar{J}$; otherwise, it is not. We adopt a naive strategy for $i \in S_j$ and $j \in \bar{J}$, distributing the residual $L_j - d_j$ equally over $|S_j|$ components. In short, the optimal solution $\boldsymbol{x}^*$ satisfies

$$[\boldsymbol{x}^*]_i = \begin{cases} b_i^k, & \text{if } i \in S_j \text{ and } j \in J_1 \cup J_2 - \bar{J}; \\ b_i^k + \dfrac{L_j - d_j}{|S_j|}, & \text{if } i \in S_j \text{ and } j \in \bar{J}; \\ 0, & \text{otherwise,} \end{cases}$$

which implies formula (22) holds. And this finishes the proof.

Appendix 7. A List of industry numbers

The self-defined industry numbers for the first-level industry classification of SWS Index are as follows:

- mining (1), media (2), electrical equipment (3), electronic (4), real estate (5),
- textile & apparel (6), non-bank finance (7), iron & steel (8), public utilities (9),
- national defense (10), chemical (11), machinery (12), computer (13),
- household appliances (14), building materials (15), building decoration (16),
- transportation (17), agriculture, forestry, animal husbandry & fisheries (18), automobile (19),
- light manufacturing (20), commercial trade (21), food & drink (22), communication (23),
- leisure service (24), medical biology (25), bank (26), non-ferrous metal (27), integrated (28).