Robust, extended goal programming with uncertainty sets: an application to a multi-objective portfolio selection problem leveraging DEA

Naeem Mohseny-Tonekabony[1] · Seyed Jafar Sadjadi[1] · Emran Mohammadi[1] ·
Mehrdad Tamiz[2] · Dylan F. Jones[3]

Abstract

This study presents a two-phase approach of Data Envelopment Analysis (DEA) and Goal Programming (GP) for portfolio selection, representing a pioneering attempt at combining these techniques within the context of portfolio selection. The approach expands on the conventional risk and return framework by incorporating additional financial factors and addressing data uncertainty, which allows for a thorough examination of portfolio outcomes while accommodating investor preferences and conservatism levels. The initial phase employs a super-efficiency DEA model to streamline asset selection by identifying suitable investment candidates based on efficiency scores, setting the stage for subsequent portfolio optimization. The second phase leverages the Extended GP (EGP) framework, which facilitates the comprehensive incorporation of investor preferences to determine the optimal weights of the efficient assets previously identified within the portfolio. Each goal is tailored to reflect specific financial factors spanning both technical and fundamental aspects. To tackle data uncertainty, robust optimization is applied. The research contributes to the robust GP (RGP) literature by analyzing new RGP variants, overcoming limitations of traditional and other uncertain GP models by incorporating uncertainty sets. Robust counterparts of the EGP models are accordingly developed using polyhedral and combined interval and polyhedral uncertainty sets, providing a flexible representation of uncertainty in financial markets. Empirical results, based on real data from the Tehran Stock Exchange comprising 779 assets, demonstrate the superiority of the proposed approach over traditional portfolio selection methods across various uncertainty settings. Additionally, a comprehensive sensitivity analysis investigates the impact of uncertainty levels on the robust EGP models. The proposed framework offers guidance to investors and fund managers through a pragmatic approach, enabling informed and robust portfolio decisions by considering efficiency, uncertainty, and extended financial factors.

Keywords Multi-objective portfolio selection · Extended goal programming · Robust optimization · Financial ratios · Data envelopment analysis

Extended author information available on the last page of the article

1 Introduction

Modern portfolio theory (MPT) emerged from the groundbreaking work of Markowitz (1952) through the introduction of the mean–variance (MV) model for asset allocation. The classic MV model primarily focuses on two criteria, i.e. return and risk. However, the intrinsic multi-dimensionality of the portfolio selection problem has been emphasized by many scholars. Empirical evidence shows that incorporating more than two factors in order to choose the best financial portfolio mitigates reliance on any single measure that might have flaws associated with it (Almeida-Filho et al., 2021; Colapinto et al., 2019; Doumpos & Zopounidis, 2014; Kim et al., 2022; Rahiminezhad Galankashi et al., 2020; Spronk et al., 2016; Steuer et al., 2007, 2008; Tamiz & Azmi, 2019; Tamiz et al., 2013; Xidonas et al., 2012). Furthermore, contemplating factors beyond risk and return is not an unsound move in a financial port-folio selection problem. It may facilitate a reduction in management distraction, as well as resulting in possible improvements to other favorable attributes (Steuer et al., 2007). Con-sequently, diversified criteria for financial portfolio selection and management align with different investment strategies which primarily rely on investor preferences. In this regard, portfolio optimization models that include more attributes in the selection of assets more real-istically represent investors' aspirations. Some of these attributes may include profitability, liquidity, systematic and non-systematic risks, financial ratios, market ratios, non-financial attributes (e.g. ethical, environmental, and social issues), skewness, and kurtosis, to name but a few (Aouni et al., 2018; Ballestero et al., 2012; Boubaker et al., 2023; Colapinto et al., 2019; Doumpos & Zopounidis, 2014; Peykani et al., 2020; Rahiminezhad Galankashi et al., 2020; Steuer et al., 2007; Wu et al., 2022; Yu & Lee, 2011).

In the context of financial portfolio selection, fundamental analysis (FA) is a method for evaluating a company's investment potential. FA involves a thorough examination of a company's financial statements in order to assess its investment-worthiness, while technical analysis (TA) relies on historical trajectories to predict future stock values (Edirisinghe & Zhang, 2007, 2008). Prior research supports the integration of FA and TA, indicating that these two techniques could be complements, rather than substitutes (Bettman et al., 2009; Contreras et al., 2012; Kuo et al., 2021; Namdari & Li, 2018). Actually, the synergy result-ing from simultaneously employing these two techniques has the potential to enhance the predicting power of firms' future financial performance, thus leading to the selection of supe-rior portfolios. In the domain of financial portfolio selection, the risk and return attributes of the investment are considered as technical factors (Kuo et al., 2021; Tamiz & Azmi, 2019). On the other hand, there are various ways to implement FA, one of which is Data Envelopment Analysis (DEA) (Abad et al., 2004; Edirisinghe & Zhang, 2007, 2008; Lim et al., 2014). DEA is a data-enabled performance evaluation technique that evaluates the relative efficiency of decision-making units (DMUs) considering several inputs and outputs (Emrouznejad & Yang, 2018; Zhu, 2022) and can be leveraged in the process of financial portfolio construction to measure assets' efficiencies, thereby specifying the best assets for investment (Edirisinghe & Zhang, 2007, 2008; Lim et al., 2014; Peykani et al., 2020, 2022). Consequently, DEA approaches can be used as a means of FA in the portfolio selection context in order to fundamentally assess a firm's investment worthiness based on compari-son to the market as a whole, rather than evaluating it in isolation, thus contributing to the identification of financially healthy firms.

Multi-criteria decision making (MCDM) approaches have the potential to enable the consideration of various factors in the portfolio selection problem, where asset screening can be combined with portfolio construction within an integrated framework (Zopounidis et al., 2015). The multi-objective programming (MOP) technique of Goal programming (GP) is a prominent and powerful MCDM approach that has been applied in various areas of financial decision-making (Andriosopoulos et al., 2019). GP reaches solutions in line with the decision maker's expressed goals and preferences in optimization problems with multiple objectives. Indeed, the GP model is based on distance functions where the undesirable positive and negative deviations between the achievement and aspiration levels of the objectives are to be minimized (Jones & Tamiz, 2010). GP is an instrumental tool to analyze portfolio selection problems and achieves reasonable solutions concerning the inclusion of the decision makers' preferences (Aouni et al., 2014; Colapinto et al., 2019). Using GP, financial decision makers (FDM) are enabled to concurrently consider various aspirational financial and non-financial factors as distinct goals, along with exclusive preferences for each of the factors in a multi-objective optimization framework. In this manner, they are able to benefit from the capabilities of this mathematical framework, including sensitivity analysis, to effectively achieve their desired portfolios. Portfolio selection problems centered only on risk and return optimization can be described as a GP model comprising of two principal goals (Tamiz & Azmi, 2019). Nonetheless, as previously mentioned, extra objectives illustrating other factors and attributes can readily be incorporated into the versatile GP model.

Another prominent issue that should be considered in any proposed approach for portfolio selection is the treatment of parameter uncertainty. The underlying uncertainty of financial markets is one of their most essential characteristics (Kim et al., 2022; Sadjadi et al., 2012). Investment decisions should be made incorporating adequate hedging against uncertainty, otherwise the resultant portfolios would not be reliable and practical for real-world applications. However, many studies assume the input data as certain. In such cases, perturbation in data values can lead to huge variances in the results to the extent that, in the worst-case, infeasible solutions may be produced. Consequently, it is necessary to develop models which meaningfully account for uncertainty while also accommodating the manifold preferences of financial decision makers.

In the optimization literature, stochastic programming (SP) and robust optimization (RO) are two of the widely-adopted techniques utilized to cope with the uncertainty. Both SP and RO seek to address the same question of building an uncertainty–immunized solution to an optimization problem with input data uncertainty. When the uncertain data are of a stochastic nature, the quality of SP-based decisions are associated with the type of underlying probability distribution. However, accurately ascertaining the underlying probability distribution is usually demanding and often necessitates a massive number of observations, rendering the SP model computationally intractable (Ben-Tal et al., 2009; Bertsimas & Sim, 2003; Ghahtarani et al., 2022; Hanks et al., 2017). Considering the aforementioned drawbacks associated with SP, RO, on the other hand, utilizes uncertainty sets instead of probability distributions to account for uncertain data and does not cause intractability. It can therefore be used as an alternative approach to address the uncertainty. RO can result in a solution that is ensured to be robust and feasible for almost all possible realizations of the uncertain parameters (Ben-Tal et al., 2009; Bertsimas & Sim, 2003). Furthermore, this method enables decision makers to make a trade-off between system reliability and economic performance. Although SP is a dynamic approach and appears to be less conservative than the worst-case

oriented and static RO, it requires a strong set of assumptions to hold, as mentioned, whereas RO is more pragmatic and hence, in the context of this paper, fits better with the pragmatic ethos of the goal programming framework. Therefore, the combination of GP and RO can result in pragmatic, flexible, and yet straightforward decision support tools for addressing multi-objective optimization problems in uncertain environments.

From the perspective of the GP methodology applied to portfolio selection problems, in this paper we follow the GP framework proposed by Tamiz et al. (2013) and Tamiz and Azmi (2019), though utilizing different factors rooted in FA, besides the conventional technical ones of risk and return, and different GP variants for the stock portfolio selection problem. Also, we extend the GP models to account for the underlying uncertainties prevalent in the financial market. A noteworthy objective in this study is to contribute to the expansion of the robust GP (RGP) literature by introducing new RGP variants that address the limitations of traditional and other uncertain GP paradigms through the incorporation of uncertainty sets. We utilize the GP problem developed in this study as a practical application case for our introduced RGP variants. In light of these advancements, we propose an innovative two-phase approach for portfolio selection, leveraging the DEA and GP techniques. The first phase involves an asset screening procedure, utilizing a super-efficiency DEA model to measure the efficiency of registered assets in the stock exchange market. This step identifies potential investment candidates based on investor-specific filtering criteria for portfolio optimization in the subsequent phase. The amount invested in each qualified asset is decided in the second phase where the portfolio is created. In this phase, a multi-factor extended goal programming (EGP) model is developed, allowing for investor-specific preferences to be considered. In this vein, each objective function represents a specific asset-related factor. Moreover, uncertainty is incorporated into the portfolio optimization phase by extending the deterministic EGP models to their robust counterparts in order to ensure model stability in the face of imprecise data. To this end, two RO approaches via a polyhedral uncertainty set (Li et al., 2011) and a combined interval and polyhedral uncertainty set (Bertsimas & Sim, 2004) are adopted and compared with each other for the following reasons: (a) RO via polyhedral uncertainty set guarantees maximum robustness against uncertainty while incurring a high cost, which entices risk averse decision makers; (b) RO via combined interval and polyhedral uncertainty set reaches a compromise between robustness and its cost, appealing to risk-seeking decision makers; (c) RO via the above-mentioned uncertainty sets doesn't contain the over-conservatism present in classic worst-case robust models, such as Soyster's (1973) approach, wherein optimality is excessively sacrificed for feasibility. Instead, in the resulting robust counterparts via the aforesaid uncertainty sets, the decision maker has complete control over the degree of conservatism for each constraint with uncertain parameters; and (d) RO via the aforementioned uncertainty sets retains the linearity of the nominal problem, in contrast to certain RO approaches (e.g. Ben-Tal & Nemirovski, 2000) which convert a linear programming problem into a nonlinear one. Using our approach, financial decision makers are able to select suitable portfolios by utilizing computationally tractable models that accommodate their specific preferences and desired levels of conservatism. Our proposed approach is validated using real data from the Tehran stock exchange (TSE) and the resulting portfolios are compared with one another, as well as with portfolios obtained from traditional benchmark models for portfolio selection, and conclusions are drawn. Figure 1 presents a schematic framework of all the steps of the proposed two-phase approach.

The remainder of this paper is organized as follows: In Sect. 2, we review the related literature on applications of DEA and GP models in financial portfolio selection followed by the robust goal programming literature and conclude by stating the research gaps and contributions. Section 3 starts with the problem description and continues with the description
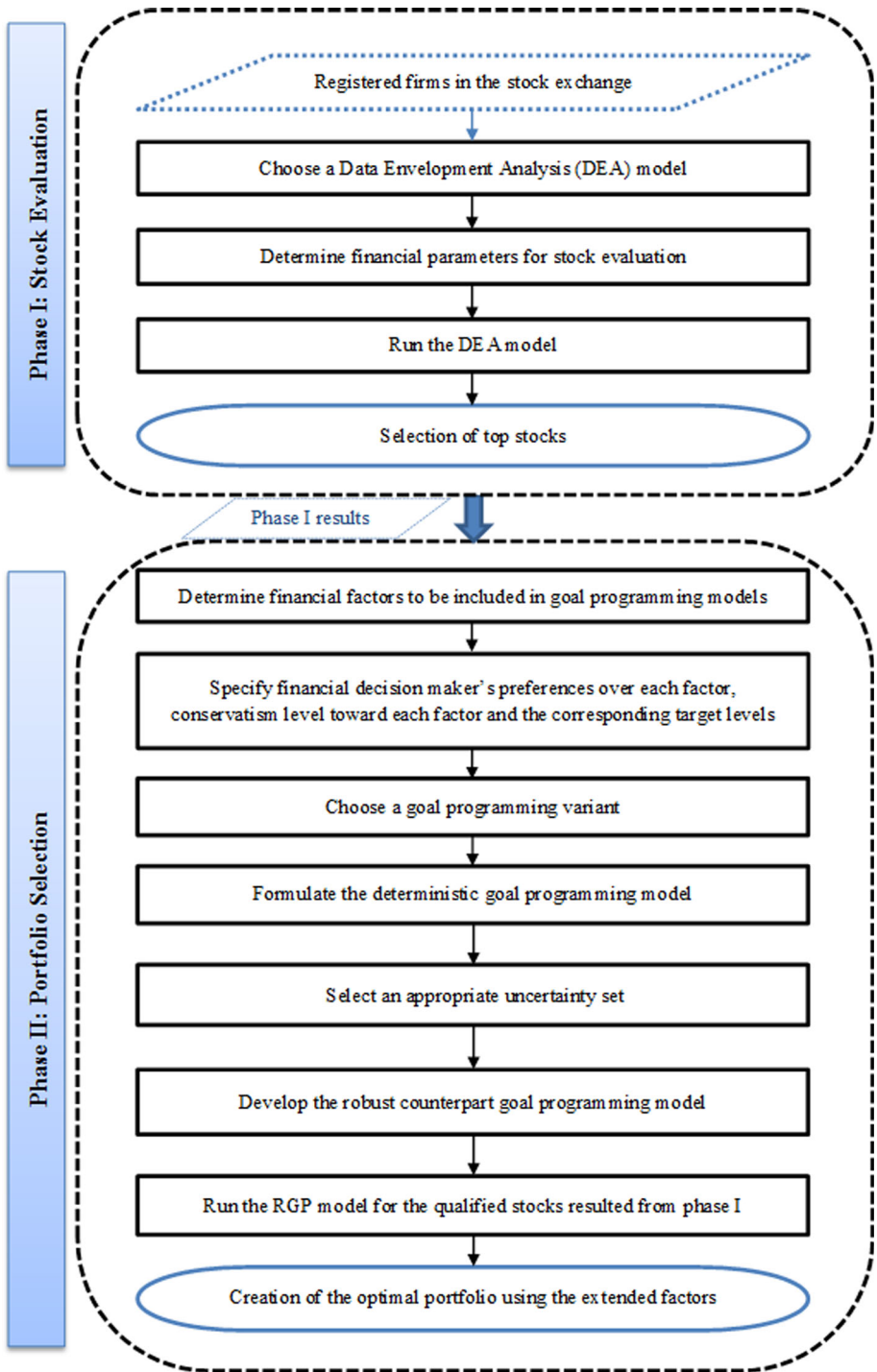
Fig. 1 The methodology of the proposed two-phase approach for the multiple criteria portfolio selection with uncertain factors

of the methodologies used in this paper, i.e. the super-efficiency DEA model, the extended factors, the EGP variant, the RO approach, and the general Robust EGP framework. The section continues with the proposed portfolio selection models by which the deterministic model is firstly introduced, and then the resulting robust counterpart GP models via the aforementioned uncertainty sets, are presented. Explanatory details about the data of the case study and various scenarios designed for the experimentations along with sensitivity analysis, results, and discussions are provided in Sect. 4. Finally, in Sect. 5 we conclude the paper and discuss directions for future research.

2 Literature review

This paper is principally related to three streams of literature: applications of DEA in financial portfolio selection, GP applied to financial portfolio selection, and robust GP. In this section, we first discuss relevant research that applies the DEA technique to the context of financial portfolio selection. Then, we describe the related literature on GP applications in financial portfolio selection. An overview of previous studies on robust GP (both methodological and applicational) is presented in subSect. 2.2. Finally, we discuss the literature gaps and research contributions.[1]

2.1 Contextual background

2.1.1 Applications of DEA in financial portfolio selection

In the portfolio selection context, there are typically two main aspects to consider: choosing which stocks to include and deciding how much to invest in each. However, many studies on portfolio selection tend to focus solely on the latter, treating it as a pure optimization problem while overlooking the importance of selecting the right assets. Traditionally, it's believed that spreading investments across a wide range of assets helps reduce risk. But in real-world investing, if a portfolio becomes too diversified, investors might end up paying higher transaction costs, and the investment strategy may not be as effective as anticipated. To address this issue, it is advisable for investors to concentrate their investments in a select number of investment-worthy stocks (Zhou et al., 2021). This approach strikes a balance between diversification and optimization, making it more practical in investment scenarios. DEA, as a non-parametric evaluation method, is among the methods that can be employed to effectively assess and identify the investment-worthy stocks.

DEA models have been used for efficiency evaluation in many fields such as finance, economics, education, athletics, and so on (Emrouznejad & Yang, 2018). The use of DEA in the portfolio selection context encompasses various purposes. DEA models have been used as a mean for efficiency estimation of portfolios post-creation. As such, various risk measures like the variance of the portfolio and CVAR have been used as inputs and the expected return of the portfolio has been used as the output in these approaches (Kuo et al., 2021; Liu et al., 2015; Tarnaud & Leleu, 2018). More recently, researchers have used DEA in the portfolio selection and creation process for performance assessment and rankings of stocks, and identifying the investment-worthy ones. In this vein, various fundamental and

---

[1] This paper can also inherently be categorized under the robust portfolio selection literature. Interested readers can refer to Ghahtarani et al. (2022) and Xidonas et al. (2020) for recent surveys on different robust portfolio selection approaches. In this paper, we are only interested in robust approaches within the GP framework.

stock market indicators have been used as inputs and outputs in classic, cross-efficiency and super-efficiency DEA models. For instance, Edirisinghe and Zhang (2007) and Edirisinghe and Zhang (2008) have similar approaches of utilizing DEA as a means of FA in the portfolio selection process. They developed generalized DEA models for evaluating and comparing firms within their industry and market context, rather than in isolation, as is the practice in common FA. The aim of their approach is to predict the future financial performance of firms, arguing about its high correlation with future stock price. They computed a relative financial strength indicator by means of analyzing a firm's financial statements and the DEA model to determine investment-worthy firms and then assigned weights to candidate stocks by using the MV framework. Pätäri et al. (2012) used the constant returns-to-scale, super-efficiency, and the cross-efficiency DEA models as a basis of selection criteria for equity portfolios, aiming at integrating the benefits of both value investing and momentum investing. Making another attempt to use DEA as an FA tool, Lim et al. (2014) developed a DEA cross-efficiency model utilizing financial indicators of listed companies and proposed an MV portfolio selection model based on two statistics of the cross-efficiency matrix. They applied their approach to stock portfolio selection within the Korean stock market and demonstrated that the selected portfolio delivered higher risk-adjusted returns compared to other benchmark portfolios. Hsu (2014) proposed an integrated procedure employing DEA, artificial bee colony (ABC) and genetic programming to address the portfolio selection problem. The input-oriented CCR model was utilized for screening stocks based on their historical financial performance and selecting the most efficient candidates for the portfolio. The inputs for the DEA model encompassed total assets, total equity, cost of sales, and operating expenses, while the outputs included net sales and net income. Subsequently, the ABC algorithm was employed to resolve a variant of the Markowitz portfolio optimization model, wherein risk was assessed using below-target semi-variance to determine investment weights. Finally, genetic programming was applied to develop a price forecasting model to guide the buy/sell decisions' timing. The approach's feasibility and effectiveness were demonstrated using Taiwanese semi-conductor stocks. Huang et al. (2015) introduced a portfolio construction approach wherein stocks, pre-screened through fundamental analysis, were subjected to a VRS DEA model for the selection of efficient stocks. Down-side risk and β-coefficient served as inputs, while return rate and Sharpe ratio were designated as outputs for the DEA model. A multi-objective decision making method based on compromise programming was employed to address the bi-objective Markowitz model and determine investment weights. They applied this methodology to Taiwan's financial market and compared it with several benchmarking mutual funds, yielding superior results in terms of return rates and Sharpe ratios. Peykani et al. (2020) proposed a two-phase approach for the portfolio selection and optimization problems in which robust DEA models were used in the first phase to measure stocks efficiency and select candidate assets for the portfolio optimization phase in the presence of uncertain data. They then applied single-objective robust mean-semi variance-liquidity and robust mean-absolute deviation-liquidity models to determine candidate investment weights and applied their methodology to the Tehran stock exchange. Amin and Hajjami (2021) developed a cross-efficiency DEA model for the portfolio selection problem and investigated the role of alternative optimal solutions in DEA models. They found that portfolio construction with lower risk and higher returns is possible when alternative optimal solutions are included in the model. In an attempt to incorporate data mining techniques into the portfolio selection and optimization scheme, Zhou et al. (2021) introduced a stock selection plan that integrated DEA with multiple data sources, then used a support vector machine (SVM) to predict the stock price movements and combined it with the stock selection scheme to construct the portfolio optimization model. The potential of DEA in combination with MODM has been

recently investigated in the context of financial instruments other than stocks. For instance, Henriques et al. (2022) present a two-step approach for portfolio modeling, utilizing DEA to identify efficient assets based on financial indicators and interval multi-objective portfolio models to determine optimal portfolio compositions according to investor preferences. Their methodology addresses computational complexities associated with portfolio optimization and uncertainty of factors in portfolio models by using DEA and interval programming. They applied their methodology to a sample of diversified Exchange Traded Funds (ETFs) operating in the US energy sector. More recent applications of DEA in stock selection can be found in the work of Hosseinzadeh et al. (2023), where the authors employed an SBM DEA model to preselect efficient assets in large-scale portfolio problems. They introduced various reward/risk criteria, and DEA input/output sets based on these criteria for asset pre-selection, subsequently evaluating the impact of preselected assets in portfolio optimization strategies. These strategies encompassed uniform investment among all the preselected assets, and strategies that individually maximize the Sharpe, stable, and Pearson ratios. The study compared DEA-based pre-selection to PCA factor models using two datasets, S&P500 components and Fama and French 100 portfolios, demonstrating superior performance of DEA pre-selection. Their approach offers advantages such as linear solvability for large-scale portfolios, identifying top assets by multiple criteria, and promising performance during systemic risk crises.

2.1.2 Goal programming in financial portfolio selection

GP is a practical method for analyzing portfolio selection problems with several incommensurable and conflicting attributes, allowing for the generation of plausible solutions that align with the FDM's preferences. Depending on the nature of the available information, the FDM is required to choose the appropriate GP variant to derive a solution that best aligns with their portfolio selection preferences (Colapinto et al., 2019). Numerous studies have been conducted on this stream of literature, and interested readers can find a recent and detailed description of works in financial portfolio management utilizing GP in Colapinto et al. (2019). To provide a comprehensive understanding of the leading keywords in GP-related research for financial portfolio selection and optimization, we conduct a bibliometric analysis using the VOS viewer software (Van Eck & Waltman, 2010), for which the result is illustrated in Fig. 2.

Figure 2 shows the density visualization of the co-occurrence of all keywords in GP-related research in financial portfolio selection. The experiment to obtain this result involves 134 records from the Web of Science Core Collection with no time period limit, and the input keywords used were "Goal Programming" and "Portfolio Selection/Management". With a threshold of at least 3 occurrences in all works, out of the 656 initial keywords, 59 were selected for the bibliometric analysis. According to Fig. 2, dense areas shown in red demonstrate the top keywords that occurred most frequently, while green areas represent keywords with lower research density in this field. For example, "uncertainty" falls into this group, indicating a lower number of works that considered uncertainty in modeling. Also, among the uncertainty modeling approaches, "fuzzy goal programming" and "stochastic goal programming" had the most occurrences, respectively. Interested readers can refer to Aouni et al. (2014) and Colapinto et al. (2019) for surveys on stochastic and fuzzy GP models in portfolio optimization problems. Another interesting point that can be deduced from the map is that a clear separation exists between the keywords like "Analytic Network Process" or "ANP" and "Goal Programming," "Portfolio Selection," "Polynomial Goal Programming," and "Fuzzy Goal Programming," which could be addressed in future research. In what follows, we will

Fig. 2 Density visualization of all the keywords with a threshold of minimum 3 occurrences

examine some relevant GP applications in financial portfolio selection and management in more detail.

Bravo et al. (2010) developed a mean value-stochastic GP model for selecting buy-and-hold efficient portfolios of funds, where they considered multiple benchmarks for returns as a more realistic approach to portfolio analysis. Stoyan and Kwon (2011) formulated a mixed-integer stochastic GP model for an integrated stock and bond portfolio problem, taking into account uncertainty in asset prices and several major trading constraints. Bilbao-Terol et al. (2012) proposed Weighted and MinMax GP models for selecting portfolios with socially responsible investment (SRI)-funds, where they measured the socially responsible performance of financial products using an index built through the application of fuzzy set theory techniques. Tamiz et al. (2013) applied three goal programming variants—Weighted, Lexicographic, and MinMax GP—to develop portfolio selection models for international mutual funds, which enable decision-makers to incorporate their preferred factors and ideal aspiration levels into the GP model for acquiring their intended portfolio. The authors selected seven factors from three groups: 'mutual funds specific factors,' 'macroeconomics factors,' and 'factors for regional preferences.' Each factor was treated as an objective in their GP models. Using data related to 20 mutual funds of equities from 10 different countries, they compared the resulting portfolios with different combinations of weights, priority levels, and target values in the GP models against each other in terms of return, risk, and the number of mutual funds selected, and demonstrated and discussed the applicability of their approach. Messaoudi et al. (2017) developed a fuzzy chance-constrained GP model to solve a financial portfolio selection problem with three attributes. In their approach, stochastic uncertainty pertains to the independent chance-constrained objectives, and the financial decision maker's preferences were considered as fuzzy values. De et al. (2018) proposed a fuzzy GP model using Werner's 'fuzzy and' hybrid operator for a portfolio selection problem with three criteria: risk, return, and liquidity. They assumed return and liquidity as fuzzy values and described

them with triangular and trapezoidal membership functions. Tamiz and Azmi (2019) utilized five FA-based factors, in addition to risk and return, for stock portfolio selection, with each factor represented as an objective in a weighted goal programming (WGP) model. Various WGP models with different combinations of target values and weights were developed. They applied their models to 30 stocks from the Dow Jones Industrial Average index, and the resulting portfolios were compared against each other, as well as against well-known benchmark models for portfolio selection from the literature. The results obtained supported the use of factors utilized besides risk and return, referred to as 'extended factors,' for addressing portfolio selection problems. Mansour et al. (2019) formulated a GP model for the financial portfolio selection problem with three criteria: risk, return, and liquidity. In this model, fuzzy returns were assumed, and investor preferences were incorporated using the concept of satisfaction functions. Deng and Yuan (2021) introduced a GP model based on fuzzy dominance for a portfolio selection problem involving fuzzy returns while simultaneously considering systematic and non-systematic risks. More recently, Bravo et al. (2022) proposed a GP approach that differs from the state-of-the-art uncertain GP models. In this approach, the variability of parameters resulting from randomness is addressed by replacing the traditional WGP achievement function with a new function that takes into account the decision maker's perception of randomness through the use of a penalty term. They applied their approach to a mean absolute deviation (MAD) portfolio selection problem and claimed that their method can effectively address the challenges arising from the lack of statistical information about random events.

2.2 Robust goal programming

While GP is extensively applied in various fields, its traditional usage has relied on deterministic values for model parameters, which is not always the case (Hanks et al., 2017). Robust goal programming (RGP) is a relatively new sub-discipline of optimization that integrates uncertainty modeling via the RO theory with traditionally deterministic GP to tackle optimization problems with multiple objectives as well as uncertain parameters.

RGP was first introduced in Kuchta's (2004) work in which parametric uncertainty is addressed through a combination of cardinality-constrained robustness and interval-based uncertainty sets within the GP technique. In this approach, the author assumed that uncertain parameters are limited to the cost coefficients in the original linear objective functions. The proposed approach allows decision-makers to find solutions for various degrees of uncertainty by changing one parameter per goal, where the worst optimal total deviation from the goals is presented. Closer to our work regarding the application context, Ghahtarani and Najafi (2013) applied Kuchta's (2004) methodology to develop a robust goal programming model for a multi-objective portfolio selection problem. Their study is the only one to date that applied the RGP methodology to portfolio selection problems modeled as a GP. The authors applied Kuchta's (2004) methodology to create the robust counterpart of Lee and Chesser's (1980) portfolio selection GP model. According to their results, as the uncertainty budget increased, solution conservatism also rose. Ghasemi Bojd and Koosha (2018) proposed a robust goal programming model for the capital budgeting problem. Their model accounts for uncertainty in cash flows and addresses multiple goals. Broadening the RGP literature, Hanks et al. (2017) extended Kuchta's (2004) work and introduced RGP models considering cardinality-constrained robustness using norm-based uncertainty sets, such as L_1 and L_2 norms, as well as strict robustness via ellipsoidal uncertainty sets. They compared the performance of their proposed models with that of Kuchta's (2004) using computational examples. The results

Table 1 An overview of RGP studies

Study	Perspective		Underlying philosophy of GP		Uncertainty set
	Methodological	Applicational	Optimization	Balance	
Kuchta (2004)	*		*		Interval + polyhedral
Ghahtarani and Najafi (2013)		*	*		Interval + polyhedral
Hanks et al. (2017)	*		*		L_1 & L_2 norms & ellipsoidal
Ghasemi Bojd and Koosha (2018)		*	*		Interval + polyhedral
Mensah and Rocca (2019)	*		*		Light robustness via budget of uncertainty & ellipsoidal
Hendalianpur et al. (2019)		*	*		Interval + polyhedral
Repetto et al. (2019)		*	*		Interval + polyhedral
Wang and Li (2019)		*	*		Interval + polyhedral
Hanks et al. (2020)		*	*		L_2-norm
Cheng et al. (2021)	*	*	*		Box ∩ generalized budget
This study	*	*	*	*	Polyhedral & interval + polyhedral

demonstrated that cardinality-constrained robustness via L_2-norm uncertainty sets outperformed cardinality-constrained robustness using interval-based uncertainty sets. However, the computational tractability of L_2-norm-based uncertainty sets still remains challenging. Leveraging the findings of the study of Hanks et al. (2017), Hanks et al. (2020) employed RGP with cardinality-constrained robustness via L_2-norm uncertainty sets to model a specific problem related to setting transportation shipping rates in a United States shipping line. Mensah and Rocca (2019), taking another step to theoretically contribute to the RGP foundation, proposed light robust goal programming models using budget of uncertainty and ellipsoidal uncertainty sets and compared their results with previous RGP models. They demonstrated that the total goal deviations of the decision-maker depend heavily on the robust quality threshold set as the trade-off between the optimality and the feasibility of the robust solution in the light RGP framework, rather than the specific uncertainty set used. Further applications of RGP can be seen in Wang and Li (2019) and Hendalianpur et al. (2019), focusing on supplier selection and order allocation in supply chain management, as well as in Repetto et al. (2019) for controlling the transfer pricing risk of a multi-national firm. A more recent application of RGP emerged in Cheng et al. (2021), where the authors developed an RGP model to optimize food distribution decisions in humanitarian relief logistics, addressing supply

and capacity uncertainties. They developed their RGP model using a refined uncertainty set, defined by the intersection of the box and generalized budget.

2.3 Research gaps and contributions

DEA and GP techniques have been independently applied to various aspects of portfolio selection and optimization in numerous studies. DEA methods have demonstrated their utility within portfolio management, serving purposes such as stock ranking in portfolio selection, estimating portfolio efficiency, and assessing market efficiency (Kuo et al., 2021; Tarnaud & Leleu, 2018). Similarly, GP has proven to be a pragmatic and versatile modeling and solution approach for addressing multi-criteria portfolio optimization problems. It enables FDMs to concurrently consider multiple incommensurable, often conflicting financial and non-financial factors as distinct goals. These factors can include unique preferences and target values, allowing FDMs to tailor their portfolios accordingly (Aouni et al., 2014; Colapinto et al., 2019). However, despite its hybridization potential (Zopounidis et al., 2015), GP has seen relatively limited integration with other approaches in the portfolio management context. Notably, Fig. 2 does not categorize DEA, indicating a scarcity of studies exploring the combination of DEA and GP in the portfolio selection context. To address this gap, we propose an integrated approach that harnesses the power of DEA and GP for portfolio selection and optimization. Furthermore, we introduce a comprehensive stock performance evaluation framework that employs the super-efficiency DEA methodology as a means of FA to identify investment-worthy stocks. In addition, selecting the appropriate GP variant is a crucial aspect of implementing the GP methodology. In our work, we utilize the EGP variant for the first time in the context of portfolio optimization to our best knowledge, as it offers superior flexibility compared to other commonly used variants. Our research explores the effectiveness of this choice in optimizing portfolios. Moreover, this study introduces another novel aspect by incorporating a set of FA-based financial factors, in addition to the TA-based factors of return and risk, as goals within the EGP models, following the concept referred to as 'extended factors' introduced by Tamiz et al. (2013). This inclusion serves to integrate and provide a more comprehensive analysis of the role of the underlying business strength of firms during the portfolio optimization phase, all viewed through the lens of multi-objective optimization. By encompassing a range of FA-based factors, the study seeks to enhance the predictive capacity of firms' future financial performance, thereby contributing to a more nuanced and effective portfolio selection process, and allowing for the identification of portfolios with greater potential for superior performance. The developed GP framework readily accommodates additional goals representing other financial or non-financial factors, providing a versatile platform for personalized portfolio optimization.

Additionally, while uncertainty within the portfolio management context is an inherent issue which needs to be accounted for, Fig. 2 as well as the relevant GP literature reveal a paucity of studies that model portfolio problems taking into account underlying uncertainties. Also, among the non-deterministic GP approaches applied to portfolio optimization, fuzzy GP and stochastic GP models have been the most frequently explored. These uncertain GP approaches have primarily focused on goal target values to accommodate uncertain conditions, which may not always align with practical GP applications, such as portfolio optimization where the primary sources of uncertainty pertain to goal function coefficients, such as return rates. To address this limitation and overcome other challenges discussed in Sect. 1, we adopt the more versatile and pragmatic RO technique to manage uncertainties within the GP model during the portfolio optimization phase. This approach also tackles

the rarity of applications of RGP models in portfolio optimization. Building upon the comprehensive review of RGP studies in Sect. 2.2 and according to Table 1, it is noteworthy that the majority of RGP models have been developed using WGP variant and the approach proposed by Bertsimas and Sim (2004) (i.e. the combined interval and polyhedral uncertainty set). Therefore, we introduce the robust counterpart of the EGP model, termed the Robust EGP (REGP) variant. Given the reasons discussed in Sect. 1, we develop the REGP models employing both pure polyhedral and combined interval and polyhedral uncertainty sets. We then compare and validate their performance using real-world data. In doing so, we aim to extend and promote the topic of RGP in line with the work of Hanks et al. (2017). RGP is an under-developed topic which deserves more attention, as it has the potential to accommodate uncertainties not only in goal function coefficients and target values but also in system constraints—a critical feature for achieving complete robustness in problems where both constraint parameters and goals are imprecisely defined. Overall, our proposed approach aims to equip FDMs with computationally tractable, adaptable, and pragmatic decision support tools for portfolio selection and optimization. The proposed approach enables FDMs to make investment decisions based on their financial preferences, including desired financial factors, preferential weights, and goal target values, while also accommodating varying levels of conservatism.

3 Methodology

3.1 Problem description

This paper proposes a multiple-criteria financial portfolio selection framework with the primary aim of maximizing the expected return of the investment as well as minimizing its risk. With the aim of incorporating the underlying business strength of firms into the portfolio selection and optimization, in the problem we consider a set of FA-based financial factors besides the technical factors of return and risk of the assets as the criteria, whilst accounting for data uncertainty. Therefore, this paper adds to the literature showing the benefits of moving beyond traditional risk-return bi-criteria approaches to portfolio selection problems with an investor who has the preference of incorporating the fundamental financial strength of firms into the portfolio selection and optimization process. In this vein, we put special focus on addressing the inherent uncertainties of financial markets, as without considering this any portfolio selection approach may be unreliable. As illustrated in Fig. 1, the proposed framework consists of 2 phases. In the first phase, through an asset screening procedure and utilizing a super-efficiency DEA model, the efficiency score of all the registered assets in the stock exchange market is measured. Accordingly, for the inputs and outputs of the DEA model, a set of financial parameters governing various aspects of the firm's operations is determined based on quantitative and qualitative information derived from financial statements of firms, as well as relevant literature and expert advice. At the end of this phase, assets that are potential investments are determined based on the investor's filtering criteria (e.g. efficient assets) as candidates for the portfolio optimization problem in the second phase. The portfolio weights of the set of qualified risky assets resulting from the first phase are decided in the second phase, where the portfolio is constructed. In this phase, a multi-factor EGP model is developed which allows for the consideration of specific investor preferences. In this regard, each goal represents a distinct asset-related financial factor, ranging from technical return and risk factors to FA-based factors derived from the first-phase DEA approach.

Further details regarding the constraints and assumptions of the portfolio optimization models will be elaborated in Sect. 3.6. Moreover, robust optimization is employed to deal with the inherent data uncertainty feature of financial markets during the portfolio optimization phase. Accordingly, the deterministic EGP models are extended to robust versions in order to ensure model stability when facing imprecise data and to strike a balance between system cost and reliability.

3.2 The super-efficiency DEA model

Various decision-making units (DMUs) can have an efficiency score of 1 in the outputs of original DEA models, such as the CCR model in Charnes et al. (1978). As a result, these models are unable to rank the DMUs precisely. In response to these flaws in traditional DEA models, Andersen and Petersen (1993) developed a novel method for ranking efficient DMUs, resulting in the notion of super-efficiency. The Andersen and Petersen (1993) model is as follows:

$$\text{Min}\,\theta_o^s$$

$$S.t.$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \overline{x}_{ij} \leq \theta_o^s \overline{x}_{io} \quad i = 1, \ldots, m$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \overline{y}_{rj} \geq \overline{y}_{ro} \quad r = 1, \ldots, s$$

$$\lambda_j \geq 0 \quad j = 1, 2, \ldots, n \tag{1}$$

where θ_o^s, $\overline{x}_{io}$, and $\overline{y}_{ro}$ are the efficiency score, input, and output values of the DMU under evaluation, respectively. The λ_j variable is the weight corresponding to the DMU_j, and the model (1) evaluates the relative efficiencies of n DMUs, each with m inputs and s outputs denoted by $\overline{x}_{1j}, \ldots, \overline{x}_{mj}$ and $\overline{y}_{1j}, \ldots, \overline{y}_{sj}$, respectively. The super-efficiency model (1) computes the efficiency score of the DMU under evaluation by removing it from constraints. It is used to rank efficient DMUs generated by the original DEA models, but it may also be applied to assess and rank all DMUs. Super-efficiency DEA models have been applied in various fields of finance for performance evaluation [see e.g. Avkiran (2011), Pätäri et al. (2012), Dutta et al. (2020), Lin and Li (2020), and Tsolas (2022) for applications in performance appraisal within banking, equity portfolio, non-banking finance companies, mutual funds, and exchange-traded funds, respectively]. Due to its superior discriminating power, the super-efficiency model (1), in its variable returns to scale (VRS) version, which includes the additional convexity constraint $\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1$, is employed in this study to fundamentally evaluate and rank all the registered assets in the stock exchange market in the first phase of the proposed methodology for the stock portfolio selection problem. The VRS model allows for the assessment of the relative efficiency of each firm by considering their specific scale of operations and capturing variations in efficiencies due to differences in the firms' financial

structures and operational characteristics. At the end of this phase, based on criteria such as a cardinality constraint or the efficiency level determined by the FDM, the top assets are selected for potential investment in the second phase.

Financial ratios are widely used and accepted in the finance and accounting literature for comparative analysis and performance evaluation purposes (Almeida-Filho et al., 2021; Wu et al., 2022). For the inputs and outputs of the DEA model, we employ a range of performance perspectives for a firm including, profitability, activity, liquidity, and leverage ratios. These four major categories of financial ratios represent the underlying financial strengths of a firm. Based on the literature (Edirisinghe & Zhang, 2007, 2008; Kuo et al., 2021; Lim et al., 2014; Peykani et al., 2022; Rahiminezhad Galankashi et al., 2020; Wu et al., 2022; Xidonas et al., 2009), expert opinions (from fund managers and financial advisors), and the Delphi method (utilizing questionnaires), we select a total of 14 critical financial parameters that encompass all of these perspectives to serve as inputs and outputs for our DEA model, as presented in Table 2. In the DEA model, parameters that follow the 'the more, the better' principle are used as outputs, while those adhering to 'the fewer, the better' principle are classified as inputs (Cook et al., 2014).

3.3 Extended factor GP models

We term the factors that we are going to use alongside the traditional TA-based risk and return factors in our proposed GP models for the portfolio optimization problem as "extended factors" (Tamiz & Azmi, 2019; Tamiz et al., 2013). In practice, the FDM may wish to oversee the construction of their portfolio by simultaneously optimizing various incommensurable and conflicting criteria, such as rate of return, risk, liquidity, dividends, number of assets, amount of short selling, social responsibility, financial ratios, investment in R&D, and so forth. These factors are often considered as constraints in the portfolio selection process. However, analysts and investors gain a comprehensive understanding of each factor's relationship with investment risk and return, as well as the trade-offs involved when each factor is treated as a distinct objective (Rahiminezhad Galankashi et al., 2020; Steuer et al., 2007, 2008). In a simple manner, all possible criteria considered as objectives can be linearly modeled (Spronk et al., 2016). Hence, we treat each factor as a distinct linearly modeled goal function within our proposed GP models tailored for the portfolio optimization phase.

It is important to note that the decision of whether to employ a specific set of factors primarily depends on the investor's preferences, strategies, and objectives. Extra goals representing other factors, financial or even non-financial, can readily be incorporated into the GP framework for portfolio optimization. In this paper, with the aim of incorporating the fundamental financial strength of firms into the portfolio optimization process and leveraging the synergy resulting from concurrently employing FA-based and TA-based factors, we define the following six factors for the multiple-criteria portfolio selection problem:

1- Return (RE); 2- Risk (RI); 3- Liquidity (LI); 4- Leverage (LE); 5- Activity (AC); 6- Profitability (PR);

Assuming i as the index of assets, the return and risk factors of asset i are assumed to be the mean and the standard deviation of the rate of returns of asset i in a given construction time period and are calculated as follows:

Table 2 Financial parameters used as inputs and outputs of the super-efficiency DEA model

Type	Parameter	Description	Perspective
Input	Current ratio	Total current assets divided by total current liabilities	Liquidity
	Quick ratio	Total current assets minus inventory divided by total current liabilities	
	Debt ratio	Total debt divided by total assets	Leverage
	Debt-equity ratio	Long-term debt divided by shareholders equity	
	Days sales outstanding	Average accounts receivable divided by the total credit sales for the period and multiplied by the number of days in the period	Activity (Asset utilization)
Output	Inventory turnover	Revenues for the period divided by average inventory	
	Working capital turnover	Revenues for the period divided by average working capital	
	Fixed asset turnover	Revenues for the period divided by average net fixed assets	
	Asset turnover	Revenues for the period divided by average total assets	
	Net profit margin	Net income divided by revenue for the period	Profitability
	Operating profit margin	Operating income divided by revenue for the period	
	Return on assets	Net income divided by the total assets	
	Return on equity	Net income divided by shareholders equity	
	Return on working capital	Net income divided by working capital	

$$r_{i,t} = \frac{p_{i,t}-p_{i,t-1}}{p_{i,t-1}} \tag{2}$$

$$RE_i = \frac{1}{T} \sum_{t=1}^{T} r_{i,t} \tag{3}$$

$$RI_i = \sqrt{\frac{\sum_{t=1}^{T} \left(r_{i,t} - RE_i\right)^2}{T-1}} \tag{4}$$

where $p_{i,t}$ and $p_{i,t-1}$ are the closing prices of asset i at the end of the time periods t and $t-1$, respectively. Also, $r_{i,t}$ is the rate of return of asset i in the time period t, T is the number

of construction time periods and RE_i and RI_i are the return and risk factors of asset i, respectively.

For the calculation of other remaining factors, the following sets and notation are introduced:

j: Index of financial parameters according to Table 2.

J_{Li}: Members of the liquidity perspective factors set according to Table 2.

J_{Le}: Members of the leverage perspective factors set according to Table 2.

J_{Ac}: Members of the activity perspective factors set according to Table 2.

J_{Pr}: Members of the profitability perspective factors set according to Table 2.

R_{ij}: The value of the jth financial parameter for the ith asset.

K_j: The normalization constant for the jth financial parameter.

$\widehat{R_{ij}}$: The normalized value of the jth financial parameter for the ith asset.

The normalized values for the financial parameters can be obtained through the following equation:

$$\widehat{R_{ij}} = \frac{R_{ij}}{K_j} \quad \forall i, j \tag{5}$$

where the normalization constant K_j is chosen to be the Euclidean mean of the values for the jth financial parameter as it is a computationally robust method among the normalization techniques (Tamiz & Azmi, 2019). Thus, we have K_j for each financial parameter j as follows:

$$K_j = \sqrt{\sum_i R_{ij}^2} \tag{6}$$

Consequently, we calculate the values of the other four factors considered in this paper beyond risk and return, called extended factors, as the arithmetic mean of the normalized values of the financial parameters belonging to each perspective as follows:

$$LI_i = \frac{\sum_{j \in J_{Li}} \widehat{R_{ij}}}{|J_{Li}|} \tag{7}$$

$$LE_i = \frac{\sum_{j \in J_{Le}} \widehat{R_{ij}}}{|J_{Le}|} \tag{8}$$

$$AC_i = \frac{\sum_{j \in J_{Ac}} \widehat{R_{ij}}}{|J_{Ac}|} \tag{9}$$

$$PR_i = \frac{\sum_{j \in J_{Pr}} \widehat{R_{ij}}}{|J_{Pr}|} \tag{10}$$

where $|J_{Li}|$, $|J_{Le}|$, $|J_{Ac}|$ and $|J_{Pr}|$ are the cardinalities of the liquidity, leverage, activity, and profitability perspective factors sets according to Table 2, respectively.

It should be noted that liquidity, leverage, activity, and profitability factors represent the fundamental financial strengths of assets in the GP models for portfolio optimization. We further note that our choice of strategic indicator-based metrics as extended factors in the GP models in this study, as described above, aligns with the existing correlations between individual ratios within each perspective according to Table 2's selection. This primarily allows for a more streamlined and manageable set of goals in the GP models (adhering to a maximum of seven objectives in an MCDM problem, according to Steuer et al. (2007)) while striving for a comprehensive assessment that encapsulates the entirety of Table 2's financial parameters. Moreover, this approach avoids losing information and critical nuances that can

Table 3 Summary of the extended factors used in this paper for the portfolio optimization problem

Factors	Objective function	Optimization type	Desired target value	Deviational variable to be penalized
Return	Z_{RE}	Maximization	T_{RE}	n_{RE}
Risk	Z_{RI}	Minimization	T_{RI}	p_{RI}
Liquidity	Z_{LI}	Minimization	T_{LI}	p_{LI}
Leverage	Z_{LE}	Minimization	T_{LE}	p_{LE}
Activity	Z_{AC}	Maximization	T_{AC}	n_{AC}
Profitability	Z_{PR}	Maximization	T_{PR}	n_{PR}

potentially result from neglecting certain ratios within a given perspective due to the existing correlations, which could be pivotal for a thorough assessment. The optimization type of the extended factors in the GP models is consistent with distinguishing the DEA inputs/outputs, as mentioned in sub-Sect. 3.2 (In the Activity perspective from Table 2, the first parameter's values were multiplied by -1 to maintain consistency with maximizing the AC factor). The extended factors utilized in this paper along with risk and return for the portfolio optimization problem, the desired type of optimization for each factor as an objective function, the desired target value of each factor, and the corresponding deviational variables to be penalized in the achievement function of the GP models are summarized in Table 3.

3.4 Extended goal programming

The extended lexicographic goal programming (ELGP) was proposed by Romero (2001) to permit a parametric analysis of trade-offs between efficiency and balance among different achievement levels of the target values. The ELGP formulation allows for the inclusion and combination of the lexicographic ordering, optimization, balancing and satisficing underlying philosophies of a single decision making entity (Jones & Tamiz, 2010). The lexicographic ordering philosophy is attainable through the precedence structure of the achievement function. The satisfying philosophy can be deduced from the set of goals. The optimizing philosophy is accomplished via the minimization of the weighted sum of deviations and the balancing philosophy is achieved through the inclusion of the maximum deviation term (D) in each priority level. Moreover, the equilibrium between optimization (efficiency) and balance (equity) can be controlled at each priority level through the parameter α which can be varied between complete concentration on optimization ($\alpha = 0$)(the WGP achievement function) and complete focus on balance ($\alpha = 1$) (the MinMax GP achievement function). Lexicographic and non-lexicographic forms of the model can be formulated for the cases of the presence and absence of a lexicographic ordering of goals, respectively. As this paper is concerned primarily with investigations of efficiency-balance trade-offs between objectives rather than prioritizing them, the non-lexicographic form of the extended goal programming model is used. The single priority level (non-lexicographic) extended goal program can be represented by the following formulation (Romero, 2004):

$$Min A = \alpha D + (1-\alpha)\left\{\sum_{i=1}^{I}(\frac{w_{n_i} n_i}{K_i} + \frac{w_{p_i} p_i}{K_i})\right\}$$

$$S.t.$$

$$\frac{w_{n_i} n_i}{k_i} + \frac{w_{p_i} p_i}{k_i} - D \leq 0 \quad i = 1, 2, \ldots, I$$

$$f_i(x) + n_i - p_i = b_i \quad i = 1, 2, \ldots, I$$

$$x \in F$$

$$n_i, p_i \geq 0 \quad i = 1, 2, \ldots, I \tag{11}$$

where A is the achievement function of the goal program, n_i is the negative deviational variable of the ith goal, w_{n_i} is the preferential weight associated with the minimization of n_i, p_i is the positive deviational variable of the ith goal, w_{p_i} is the preferential weight associated with the minimization of p_i, K_i is the normalization constant associated with the ith goal, b_i represents the target level for the ith goal, $f_i(x)$ is the ith objective function, F represents the feasible region of the original multi-objective problem, I is the number of goals and finally parameter α provides a trade-off between minimization of the weighted sum of unwanted deviation variables and minimization of the maximum deviation from the target values.

Despite its flexibility and advantages over other more commonly used GP variants, the EGP variant is a relatively less used variant in the literature in solving multi-objective optimization problems. To date, some of the data-driven applications of the EGP in the literature include integrated lot-sizing and cutting stock (Oliveira et al., 2021), allocation of medical robotic devices to treatment centres (Jones et el. 2022), supply chain planning in the silk industry (Jatuphatwarodom et al., 2018), and energy and asset management in an industrial microgrid (Choobineh & Mohagheghi, 2016). In this research, we explore the effectiveness of the EGP variant in the financial portfolio selection context under both deterministic and uncertain conditions.

### 3.5 Robust optimization framework

Unlike stochastic optimization, robust optimization utilizes uncertainty sets instead of probability distributions to account for uncertain data. In our study, two RO approaches via polyhedral uncertainty set (Li et al., 2011) and combined interval and polyhedral uncertainty set (Bertsimas & Sim, 2004) are adopted to develop and compare the robust counterparts of the deterministic extended goal programming model, due to the advantages discussed in Sect. 1. To present these two RO approaches, let us consider a linear mathematical programming model as:

$$Max \sum_{j} c_j x_j$$

$$S.t.$$

$$\sum_{j} \widetilde{a}_{ij} x_j \leq b_i \quad \forall i$$

$$x \in X \tag{12}$$

where coefficients $\widetilde{a}_{ij}$ are subject to uncertainty. Let J_i denote the set of uncertain coefficients in the i_{th} constraint. Each uncertain parameter $\widetilde{a}_{ij}$ with a symmetric and bounded distribution in interval $\left[a_{ij} - \widehat{a}_{ij}, a_{ij} + \widehat{a}_{ij}\right]$ can be demonstrated as:

$$\widetilde{a}_{ij} = a_{ij} + \zeta_{ij}\widehat{a}_{ij} \quad \forall j \in J_i \tag{13}$$

Fig. 3 Illustration of polyhedral uncertainty set
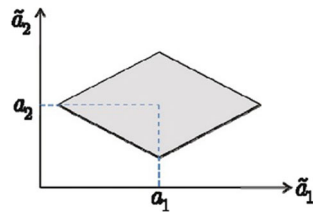
where a_{ij} is the nominal value, $\widehat{a}_{ij}$ represents the perturbation magnitude of each uncertain parameter $\widetilde{a}_{ij}$, and ζ_{ij} is an independent random variable taking values in range $[-1, 1]$. According to this definition, Ben-Tal and Nemirovski (1998) show that the set induced robust counterpart formulation of the original i_{th} constraint can be expressed as:

$$\sum_j a_{ij} x_j + \left[\max_{\zeta \in U} \left\{ \sum_{j \in J_i} \zeta_{ij} \widehat{a}_{ij} x_j \right\} \right] \le b_i \quad \forall i \tag{14}$$

It is obvious that the robust formulation (14) is closely linked to the uncertainty set U which is defined using an arbitrary norm on vector spaces. A polyhedral uncertainty set (as shown in Fig. 3 for a 2-dimensional uncertain parameter space) is employed at first to develop the robust counterpart models. If the uncertainty set U is chosen to be a polyhedral uncertainty set defined by the L^1 norm of the random variable ζ, it can be expressed as:

$$U_1 = \{\zeta | \|\zeta\|_1 \le \Gamma\} = \left\{ \zeta | \sum_{j \in J_i} |\zeta_j| \le \Gamma, \forall j \right\} \tag{15}$$

where Γ is an adjustable parameter which controls the size of the uncertainty set. In this case, the robust counterpart formulation (14) will be equivalent to the following set of constraints (Li et al., 2011):

$$\sum_j a_{ij} x_j + \Gamma_i \lambda_i \le b_i \quad \forall i$$

$$\lambda_i \ge \widehat{a}_{ij} y_j \quad \forall i, j \in j_i$$

$$-y_j \le x_j \le y_j \quad \forall j$$

$$\lambda_i \ge 0 \quad \forall i$$

$$y_j \ge 0 \quad \forall j \tag{16}$$

where λ_i is a new auxiliary variable and Γ_i, which is not necessarily an integer, can take value from interval $[0, |J_i|]$, where $|J_i|$ denotes the cardinality of set J_i. Parameter Γ_i aims to

Fig. 4 Illustration of combined interval and polyhedral uncertainty set

$0 < \Gamma < 1$ $\Gamma = 1$ $1 < \Gamma < |J_i|$ $\Gamma = |J_i|$
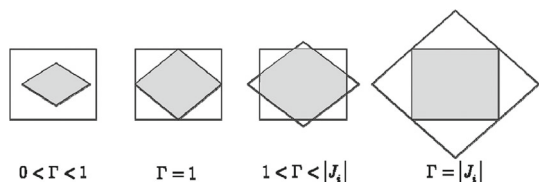
allow $\lfloor \Gamma_i \rfloor$ coefficients of row i to take their worst-case value and is used to adjust the trade-off between the robustness of the model and the conservatism level of the solution. Also, it should be noted that there will be no need to utilize the variable y_j in the case that $x_j \geq 0, \forall j$.

Subsequently, the behavior of the robust counterpart models is investigated via a combined interval and polyhedral uncertainty set (Fig. 4). This uncertainty set defined by the intersection of L^∞ and L^1 norms of random variable ζ is described as follows (Li et al., 2011):

$$U_{1\cap\infty} = \left\{ \zeta \mid \sum_{j \in J_i} |\zeta_j| \leq \Gamma, |\zeta_j| \leq 1, \forall j \right\} \tag{17}$$

When the uncertainty set U is defined as the Eq. (17), the equivalent formulation of the robust counterpart (14) is exactly similar to the robust approach proposed by Bertsimas and Sim (2004), as follows:

$$\sum_j a_{ij}x_j + \Gamma_i \lambda_i + \sum_{j \in J_i} \omega_{ij} \leq b_i \quad \forall i$$

$$\lambda_i + \omega_{ij} \geq \widehat{a}_{ij} y_j \quad \forall i, j \in j_i$$

$$-y_j \leq x_j \leq y_j \quad \forall j$$

$$\omega_{ij} \geq 0 \quad \forall i, j \in J_i$$

$$\lambda_i \geq 0 \quad \forall i$$

$$y_j \geq 0 \quad \forall j \tag{18}$$

where λ_i and ω_{ij} are new auxiliary variables and similar to the robust counterpart formulation via the polyhedral uncertainty set, a parameter $\Gamma_i \in [0, |J_i|]$, called the uncertainty budget, is introduced for each constraint i to offer a mechanism to control the deterioration effect on the objective function value against the probability of the constraint's violation. It is worth noting that in the aforesaid robust counterpart formulations, when Γ_i is set equal to zero, the constraints are equivalent to that of the nominal problem. Similarly, when Γ_i is set equal to $|J_i|$, the robust model acts as conservative as in the robust formulation of Soyster's (1973).

In the above-mentioned robust counterpart approaches, the decision maker can adjust the conservatism level of constraint i by changing the value of parameter Γ_i. Since the parameter Γ_i is an input of the robust formulation, the decision-maker can choose it according to his/her risk aversion level. The larger Γ_i is, the more risk-averse the decision maker is. If up to Γ_i number of uncertain coefficients perturb from their nominal values, the robust solution will remain feasible for every change. However, the upper bound for the probability of constraint violation can be calculated as $\exp(-\Gamma_i^2/2|J_i|)$ (Li & Floudas, 2012). Similar steps can also be followed to utilize the aforementioned robust approaches for tackling uncertainty in the right hand side (RHS) parameters.

3.5.1 General framework of the robust extended goal programming (REGP)

Assuming $f_i(x)$ as linear objective functions and goals of the minimization type in the general non-lexicographic extended goal programming formulation given in (11), the general robust

extended goal programming framework using the polyhedral uncertainty set is formulated as follows:

$$Min A = \alpha D + (1 - \alpha) \left\{ \sum_{i=1}^{I} \left(\frac{w_{p_i} p_i}{K_i} \right) \right\}$$

$$S.t.$$

$$\frac{w_{p_i} p_i}{k_i} - D \leq 0 \quad i = 1, 2, \ldots, I$$

$$\sum_{j} c_{ij} x_j + \Gamma_i \lambda_i + n_i - p_i = b_i \quad i = 1, 2, \ldots, I$$

$$\lambda_i \geq \widehat{c}_{ij} x_j \quad i = 1, 2, \ldots, I, j \in j_i$$

$$x \in F$$

$$x_j, n_i, p_i, \lambda_i \geq 0 \quad i = 1, 2, \ldots, I, \forall j \qquad (19)$$

where c_{ij} represents the nominal values of the uncertain objective function coefficients, $\widehat{c}_{ij}$ denotes the perturbation magnitude corresponding to the uncertain coefficient c_{ij}, Γ_i is the conservatism level corresponding to the ith goal constraint, λ_i is the auxiliary variable of the robust counterpart model associated with the ith goal constraint, and the significance of all other notations and constraints remains the same as defined in formulation (11).

Under the same assumptions, the general robust extended goal programming framework utilizing the combined interval and polyhedral uncertainty set is represented as follows:

$$Min A = \alpha D + (1 - \alpha) \left\{ \sum_{i=1}^{I} \left(\frac{w_{p_i} p_i}{K_i} \right) \right\}$$

$$S.t.$$

$$\frac{w_{p_i} p_i}{k_i} - D \leq 0 \quad i = 1, 2, \ldots, I$$

$$\sum_{j} c_{ij} x_j + \Gamma_i \lambda_i + \sum_{j \in J_i} \omega_{ij} + n_i - p_i = b_i \quad i = 1, 2, \ldots, I$$

$$\lambda_i + \omega_{ij} \geq \widehat{c}_{ij} x_j \quad i = 1, 2, \ldots, I, j \in j_i$$

$$x \in F$$

$$\omega_{ij} \geq 0 \quad i = 1, 2, \ldots, I, j \in j_i$$

$$x_j, n_i, p_i, \lambda_i \geq 0 \quad i = 1, 2, \ldots, I, \forall j \qquad (20)$$

where ω_{ij} is the auxiliary variable of the robust counterpart model associated with the ith goal constraint and decision variable x_j, and all other notations and constraints are consistent with those in the REGP model described earlier via the polyhedral uncertainty set.

3.6 The proposed portfolio selection models

In the second phase of the proposed methodology for the stock portfolio selection problem, the amount to be invested in each qualifying stock resulting from the first phase is decided and finally the desired portfolio will be constructed. To this end, a multi-factor extended goal programming model containing all the factors introduced in Sect. 3.3 is developed. Then, taking into account the underlying uncertainties in different factors, the robust counterparts of the EGP models are developed utilizing the two different uncertainty sets mentioned in Sect. 1. It is worth noting that in order to make the models general and comparable with classic benchmark models, no additional constraints other than the budget constraint were considered in the portfolio selection models. Naturally, such constraints like cardinality constraint, transaction costs constraint and many other market-related constraints could be straightforwardly added to the models according to the investors' preferences. Furthermore, no short-selling is allowed to make the whole problem more realistic.

Prior to setting forth the various formulations we investigate, it is necessary to introduce the associated sets, parameters, and decision variables. The notations are defined as follows:

3.6.1 Indices

i Set of assets, $i = 1, 2, \ldots, N$
q Set of factors, $q = RE, RI, LI, LE, AC, PR$

3.6.2 Parameters

RE_i Nominal value corresponding to the return factor of the ith asset.
RI_i Nominal value corresponding to the risk factor of the ith asset.
LI_i Nominal value corresponding to the liquidity factor of the ith asset.
LE_i Nominal value corresponding to the leverage factor of the ith asset.
AC_i Nominal value corresponding to the activity factor of the ith asset.
PR_i Nominal value corresponding to the profitability factor of the ith asset
v_q Weight attributed to each negative deviational variable, $\forall q$
w_q Weight attributed to each positive deviational variable, $\forall q$
T_q Target values for goals, $\forall q$
k_q Normalization constant for deviational variables, $\forall q$
α Parameter for controlling the trade-off between efficiency and equity in the EGP achievement function
Γ_q Level of conservatism (budget of uncertainty), $\forall q$
Δ_q Maximum deviation corresponding to each uncertain factor (percentage of each factor), $\forall q$

3.6.3 Decision variables

x_i Proportion of funds invested in the ith asset.
n_q Negative deviational variables, $\forall q$
p_q Positive deviational variables, $\forall q$
D Maximum weighted deviation from amongst the set of unwanted deviations.
ω_i^q Penalty associated with each factor and decision variable x_i (auxiliary variable of the robust counterpart model), $\forall q$

λ_q Scalar value that takes into account ω_i (auxiliary variable of the robust counterpart model), $\forall q$

3.6.4 The proposed deterministic extended goal programming model for portfolio selection

Considering all the factors previously described in Table 3, the proposed multi-factor deterministic extended goal programming model for the portfolio selection problem has the following general structure:

$$MinZ = \alpha D + (1-\alpha)\left(\frac{v_{RE}}{k_{RE}}n_{RE} + \frac{v_{AC}}{k_{AC}}n_{AC} + \frac{v_{PR}}{k_{PR}}n_{PR} + \frac{w_{RI}}{k_{RI}}p_{RI} + \frac{w_{LI}}{k_{LI}}p_{LI} + \frac{w_{LE}}{k_{LE}}p_{LE}\right) \tag{21}$$

$$S.t.$$

$$\frac{v_{RE}}{k_{RE}}n_{RE} - D \leq 0 \tag{22}$$

$$\frac{w_{RI}}{k_{RI}}p_{RI} - D \leq 0 \tag{23}$$

$$\frac{w_{LI}}{k_{LI}}p_{LI} - D \leq 0 \tag{24}$$

$$\frac{w_{LE}}{k_{LE}}p_{LE} - D \leq 0 \tag{25}$$

$$\frac{v_{AC}}{k_{AC}}n_{AC} - D \leq 0 \tag{26}$$

$$\frac{v_{PR}}{k_{PR}}n_{PR} - D \leq 0 \tag{27}$$

$$\sum_{i=1}^{n} RE_i x_i + n_{RE} - p_{RE} = T_{RE} \tag{28}$$

$$\sum_{i=1}^{n} RI_i x_i + n_{RI} - p_{RI} = T_{RI} \tag{29}$$

$$\sum_{i=1}^{n} LI_i x_i + n_{LI} - p_{LI} = T_{LI} \tag{30}$$

$$\sum_{i=1}^{n} LE_i x_i + n_{LE} - p_{LE} = T_{LE} \tag{31}$$

$$\sum_{i=1}^{n} AC_i x_i + n_{AC} - p_{AC} = T_{AC} \tag{32}$$

$$\sum_{i=1}^{n} PR_i x_i + n_{PR} - p_{PR} = T_{PR} \tag{33}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{34}$$

$$x_i \geq 0 \quad \forall i \tag{35}$$

$$n_{RE}, p_{RE}, n_{RI}, p_{RI}, n_{LI}, p_{LI}, n_{LE}, p_{LE}, n_{AC}, p_{AC}, n_{PR}, p_{PR} \geq 0 \tag{36}$$

where the achievement function (21) minimizes the unwanted deviations according to the EGP philosophy, constraints (22)–(27) determine the maximal weighted, normalized

deviation from amongst the set of unwanted deviations, constraints (28)–(33) calculate the positive and negative deviations from the desired target value for each goal for a given solution x_i, $\forall i$, constraint (34) is the budget constraint in the portfolio selection problem, constraints (35) indicates that no short-selling is allowed, and constraint set (36) ensures the non-negativity of all the deviational variables.

3.6.5 The proposed robust extended goal programming models (REGP) for portfolio selection

This sub-section intends to extend the deterministic goal programming models to robust counterpart models in which different financial factors are considered uncertain. This is done via considering two different uncertainty sets previously described. The uncertain parameters are modeled as random variables which take values according to symmetric distributions with means equal to the nominal values. For instance, RE_i belongs to the interval $\left[RE_i - \widehat{RE}_i, RE_i + \widehat{RE}_i \right]$, where RE_i denotes the nominal value and $\widehat{RE}_i$ denotes the perturbation magnitude. Also, the perturbation magnitude corresponding to each factor is considered as a percentage (Δ) of the nominal value of each factor. E.g., $\widehat{RE}_i = \Delta \times RE_i$, where Δ can be different for each of the factors.

According to the aforementioned assumptions and the robust formulations expressed previously, the corresponding robust counterpart of the model (21)-(36) via the polyhedral uncertainty set can be described as follows:

$$MinZ = \alpha D + (1 - \alpha)\left(\frac{v_{RE}}{k_{RE}} n_{RE} + \frac{v_{AC}}{k_{AC}} n_{AC} + \frac{v_{PR}}{k_{PR}} n_{PR} + \frac{w_{RI}}{k_{RI}} p_{RI} + \frac{w_{LI}}{k_{LI}} p_{LI} + \frac{w_{LE}}{k_{LE}} p_{LE} \right) \tag{37}$$

$$S.t.$$

$$\sum_{i=1}^{n} RE_i x_i - \Gamma_{RE}\lambda_{RE} + n_{RE} - p_{RE} = T_{RE} \tag{38}$$

$$\sum_{i=1}^{n} RI_i x_i + \Gamma_{RI}\lambda_{RI} + n_{RI} - p_{RI} = T_{RI} \tag{39}$$

$$\sum_{i=1}^{n} LI_i x_i + \Gamma_{LI}\lambda_{LI} + n_{LI} - p_{LI} = T_{LI} \tag{40}$$

$$\sum_{i=1}^{n} LE_i x_i + \Gamma_{LE}\lambda_{LE} + n_{LE} - p_{LE} = T_{LE} \tag{41}$$

$$\sum_{i=1}^{n} AC_i x_i - \Gamma_{AC}\lambda_{AC} + n_{AC} - p_{AC} = T_{AC} \tag{42}$$

$$\sum_{i=1}^{n} PR_i x_i - \Gamma_{PR}\lambda_{PR} + n_{PR} - p_{PR} = T_{PR} \tag{43}$$

$$\lambda_{RE} \geq \widehat{RE}_i x_i \quad \forall i \tag{44}$$

$$\lambda_{RI} \geq \widehat{RI}_i x_i \quad \forall i \tag{45}$$

$$\lambda_{LI} \geq \widehat{LI}_i x_i \quad \forall i \tag{46}$$

$$\lambda_{LE} \geq \widehat{LE}_i x_i \quad \forall i \tag{47}$$

$$\lambda_{AC} \geq \widehat{AC}_i x_i \quad \forall i \tag{48}$$

$$\lambda_{PR} \geq \widehat{PR}_i x_i \quad \forall i \tag{49}$$

$$\lambda_{RE}, \lambda_{RI}, \lambda_{LI}, \lambda_{LE}, \lambda_{AC}, \lambda_{PR} \geq 0 \tag{50}$$

Constraints (22) to (27).

Constraints (34) to (36).

where the significance of all notations and constraints remains the same as that defined in model (21)-(36), except that the goal constraints (28)-(33) contain an additional term, $\pm\Gamma_q\lambda_q, \forall q$, with respect to the optimization type, which accounts for the maximum variability they induce, given a conservatism level Γ_q, and which is bounded for each combination of decision variable and goal via the additional constraints set (44)-(49). Finally, constraint set (50) enforces non-negativity restrictions on auxiliary variables $\lambda_q, \forall q$.

The robust counterpart formulation of the model (21)-(36) via the combined interval and polyhedral uncertainty set (Bertsimas and Sim's (2004) approach) can be derived in a parallel way and is represented as follows:

$$MinZ = \alpha D + (1-\alpha)\left(\frac{v_{RE}}{k_{RE}}n_{RE} + \frac{v_{AC}}{k_{AC}}n_{AC} + \frac{v_{PR}}{k_{PR}}n_{PR} + \frac{w_{RI}}{k_{RI}}p_{RI} + \frac{w_{LI}}{k_{LI}}p_{LI} + \frac{w_{LE}}{k_{LE}}p_{LE}\right) \tag{51}$$

$$S.t.$$

$$\sum_{i=1}^{n} RE_i x_i - \Gamma_{RE}\lambda_{RE} - \sum_{i=1}^{n} \omega_i^{RE} + n_{RE} - p_{RE} = T_{RE} \tag{52}$$

$$\sum_{i=1}^{n} RI_i x_i + \Gamma_{RI}\lambda_{RI} + \sum_{i=1}^{n} \omega_i^{RI} + n_{RI} - p_{RI} = T_{RI} \tag{53}$$

$$\sum_{i=1}^{n} LI_i x_i + \Gamma_{LI}\lambda_{LI} + \sum_{i=1}^{n} \omega_i^{LI} + n_{LI} - p_{LI} = T_{LI} \tag{54}$$

$$\sum_{i=1}^{n} LE_i x_i + \Gamma_{LE}\lambda_{LE} + \sum_{i=1}^{n} \omega_i^{LE} + n_{LE} - p_{LE} = T_{LE} \tag{55}$$

$$\sum_{i=1}^{n} AC_i x_i - \Gamma_{AC}\lambda_{AC} - \sum_{i=1}^{n} \omega_i^{AC} + n_{AC} - p_{AC} = T_{AC} \tag{56}$$

$$\sum_{i=1}^{n} PR_i x_i - \Gamma_{PR}\lambda_{PR} - \sum_{i=1}^{n} \omega_i^{PR} + n_{PR} - p_{PR} = T_{PR} \tag{57}$$

$$\lambda_{RE} + \omega_i^{RE} \geq \widehat{RE}_i x_i \quad \forall i \tag{58}$$

$$\lambda_{RI} + \omega_i^{RI} \geq \widehat{RI}_i x_i \quad \forall i \tag{59}$$

$$\lambda_{LI} + \omega_i^{LI} \geq \widehat{LI}_i x_i \quad \forall i \tag{60}$$

$$\lambda_{LE} + \omega_i^{LE} \geq \widehat{LE}_i x_i \quad \forall i \tag{61}$$

$$\lambda_{AC} + \omega_i^{AC} \geq \widehat{AC}_i x_i \quad \forall i \tag{62}$$

$$\lambda_{PR} + \omega_i^{PR} \geq \widehat{PR}_i x_i \quad \forall i \tag{63}$$

$$\omega_i^{RE}, \omega_i^{RI}, \omega_i^{LI}, \omega_i^{LE}, \omega_i^{AC}, \omega_i^{PR} \geq 0 \quad \forall i \tag{64}$$

Constraints (22) to (27).

Constraints (34) to (36), (50).

where all notations and constraints are in common with the REGP model via the polyhedral uncertainty set described first, except that the term for the maximum variability each goal constraint defined in (28)-(33) induces will include $\pm \Gamma_q \lambda_q$ and $\pm \sum_{i=1}^n \omega_i^q$, $\forall q$, with respect to the optimization type, leading to the subsequent changes to the constraints set (44)-(49) which result in the new set of constraints (58)-(63), as well as the additional non-negativity restrictions on auxiliary variables ω_i^q, $\forall i, q$, as defined by constraint set (64).

4 Experiments and results

4.1 Data description

The Tehran Stock Exchange (TSE) was chosen as a favorable real-world case study for the problem investigated in this research. With over half a century of history, TSE is a major stock market in Iran and the Middle East. Recent fluctuations in market dynamics have heightened volatility within the TSE. These fluctuations often bring uncertainties in key financial metrics of firms that inform investor portfolio decisions. Therefore, deploying robust and resilient investment strategies that shield investors from such uncertainties becomes crucial for investing in this market, making TSE a suitable testing ground for the proposed robust investment approaches in this study. [see e.g. Peykani et al. (2020), (2022), and Rahiminezhad Galankashi et al. (2020) for some other relevant studies on TSE]. Moreover, the inherent flexibility of the proposed approach suggests confidence in its adaptability to navigate various market conditions effectively, enhancing its potential applicability across a broader spectrum of markets. For the first phase, a total of 779 firms listed on the TSE were investigated in order to be evaluated and ranked by the DEA model. The data required for the inputs and outputs of the DEA model according to Table 2 were extracted from the publicly-available financial statements of the firms, all of which are as of June 2019. The experiments for setting up the portfolios in the second phase used a constructing period of 12 months, from November 2018 to November 2019. Besides, the out of sample performance of the resultant portfolios was evaluated using a testing period of 6 months, from December 2019 to May 2020. Also, the computational experiments of this study regarding the optimization models were performed using GAMS software with the CPLEX solver.

4.2 Phase I results

Financial ratios associated with a firm's financial performance, especially those within the same perspective, exhibit correlations (Chen, 2008; Wu et al., 2022), which could be a problem for DEA computations, particularly considering the parameter selection outlined in Table 2 in this study. Neglecting certain ratios within a given financial perspective due to their correlations, on the other hand, can potentially result in a loss of information and critical nuances that could be pivotal for a comprehensive assessment. In this study, we have adopted an aggregation approach that takes into account both the correlations between ratios and their significance in representing specific financial perspectives to streamline the analysis of financial performance indicators within the context of DEA. This involves transforming

the diverse set of financial ratios into a concise set of composite variables, each labeled to convey its intended perspective. Specifically, we have created: (1) Input 1, termed "Liquidity Score," as a weighted sum of the normalized values of "Current Ratio" and "Quick Ratio" to represent the liquidity perspective; (2) Input 2, denoted as "Leverage Score," is derived from the weighted sum of the normalized values of "Debt Ratio" and "Debt-to-Equity Ratio" reflecting the leverage perspective; (3) Input 3, "Days Sales Outstanding (DSO)", is treated as a standalone input, as 'the fewer, the better' is the preferred criterion for it, in contrast to 'the more, the better' as the ideal principle applied to all other ratios within the activity (asset utilization) perspective; (4) Output 1, labeled as "Asset Utilization Score," is a weighted sum of the normalized values of "Inventory Turnover," "Working Capital Turnover," "Fixed Asset Turnover," and "Asset Turnover"; and (5) Output 2, termed "Profitability Score," encapsulates the normalized values of "Net Profit Margin," "Operating Profit Margin," "Return on Assets," "Return on Equity," and "Return on Working Capital." The choice of weights should reflect the preferences and objectives of the FDMs and can be determined with the help of experts. To ensure a balanced analysis, we employ an equal weighting scheme across these aggregated variables. It should be noted that the chosen aggregation structure in the DEA model in this study aligns with the investor-oriented objective of identifying fundamentally efficient firms for investment, emphasizing a holistic view of financial health over specific areas of strength or weakness. Further, with the aim of meeting the condition of homogeneity of the DMUs, the values used in the DEA computations for the "current" and "quick" ratios are the distance to the industry average for each asset, thereby eliminating the role of industry type. In addition to the weighted sum approach utilized in this study, future research could explore complementary methods, such as clustering techniques as suggested by Wu et al. (2022), for feature reduction in this context.

After running the super-efficiency DEA model with the specifications stated above for all 779 assets registered on the TSE, the efficiency score for each asset is obtained and the assets are ranked based on the scores. The efficiency characteristic of each asset, as determined by the DEA model results, serves as the screening philosophy. Therefore, the selection criterion is chosen to be assets with efficiency scores strictly greater than 1. In this regard, the 60 top assets are chosen. However, due to the lack of historical data necessary for calculating risk and return factors because of permanent deregistration from the stock exchange market during the construction period, we had to exclude 20 assets from the list of the top 60. As a result, we identified and selected the remaining 40 assets as qualified candidates for the portfolio optimization problem in the second phase. Table 6 includes the efficiency scores and ranking of these 40 qualified assets resulting from phase I.

4.2.1 DEA model validation and sensitivity analysis

A sensitivity analysis of the super-efficiency DEA model was conducted to assess the internal and external validity of the Phase I findings (Parkin & Hollingsworth, 1997; Habib & Kayani, 2023). The internal validity test involves the elimination of input and output variables to examine their effects on the DEA efficiency scores. In the current study, the input and output variables of the basic DEA model were removed from the efficiency model sequentially. To gain insights into the distribution of the super-efficiency scores derived from the original DEA model, the Anderson–Darling normality test is conducted (Fig. 5), which reveals that the data are not normally distributed. The test was also conducted for the modified DEA models, which similarly showed that the modified super-efficiency scores did not follow a normal distribution. Hence, a Mann–Whitney U test and the Kruskal–Wallis test are conducted to compare the efficiency scores resulting from the modified DEA models with the original
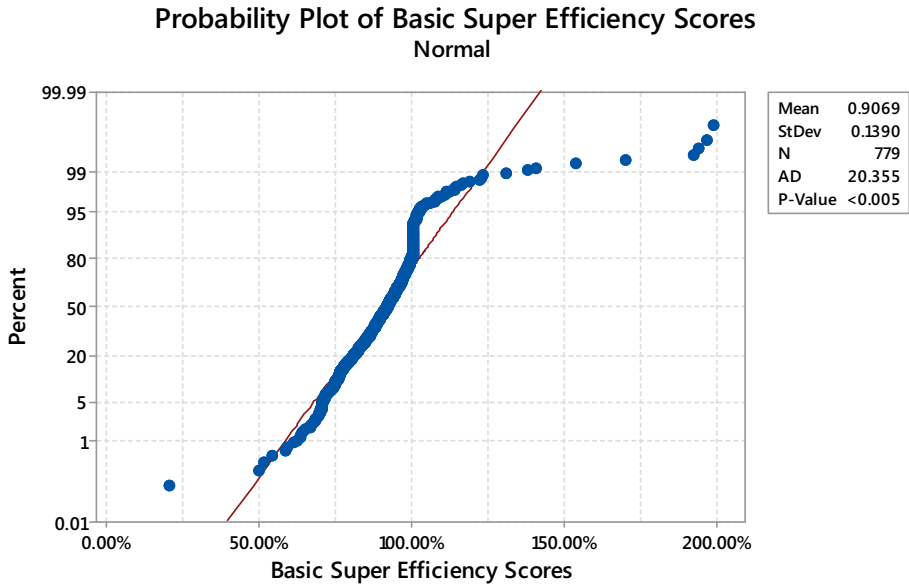
Fig. 5 Anderson–Darling normality test of the basic super-efficiency scores

scores to determine whether the removal of the variables resulted in a statistically significant difference in efficiency scores (at a significance level of 0.05). Additionally, the Spearman rank correlation coefficients were calculated to determine whether the rankings of the firms changed within the DEA models. The results of the super-efficiency DEA models with modified specifications are presented in Table 4. As indicated in Table 4, the removal of either input 2 (Leverage Score) or input 3 (DSO) had a significant impact on the model results in terms of the general distribution of the efficiency scores, as evident by the notable drop in average efficiency. However, it had a lesser effect on the rankings of firms, as evidenced by the high and significant Spearman rank correlation coefficients. This outcome is expected, given that either input measure two distinct resource categories. Consequently, removing either one would result in significant information loss. On the output side, none of the outputs (Asset Utilization and Profitability scores) appeared to significantly alter the model results when individually removed, as evidenced by the results of the Mann–Whitney U-test and the significant and high coefficients of the Spearman correlation. Such findings are in line with expectations, as the model is input-oriented.

To assess the external validity of the super-efficiency DEA model, a longitudinal analysis is conducted aimed at examining the consistency of the results over time. The original super-efficiency model was re-applied using data from a year before and a year after the original data set, for which the results are presented in Table 5. Subsequently, the efficiency scores were compared to the original results. The Mann–Whitney U-test revealed no statistically significant difference in the distribution of efficiency scores across the study years, as all P-values exceeded the significance level of 0.05. Furthermore, The Kruskal–Wallis test also supported the results of Mann–Whitney U-test, indicating no statistically significant difference in the efficiency score distribution over the years (P-value $= 0.480$). Additionally, The Spearman rank correlation coefficients between each year were also high and significant. This suggests that neither the overall distribution of efficiency scores nor the rankings of

Table 4 Sensitivity analysis of the super-efficiency DEA model

Variable removed	Average efficiency score	P-value(Mann–Whitney U test)	Spearman rank correlation coefficient (significance)
None	0.907	-	-
Input 2	0.456	0.000	0.844 (0.000)
Input 3	0.761	0.000	0.920 (0.000)
Output 1	0.690	0.125	0.972 (0.000)
Output 2	0.819	0.289	0.980 (0.000)
P-value(Kruskal–Wallis test)		0.000	

Table 5 Longitudinal analysis of the super-efficiency DEA model

Year	P-value(Mann–Whitney U test)	Spearman rank correlation coefficient (significance)
2018–2019	0.961	0.657 (0.000)
2018–2020	0.341	0.432 (0.000)
2019–2020	0.254	0.484 (0.000)
P-value (Kruskal–Wallis test)	0.480	

assets displayed substantial variation from one year to the next. These results underscore the coherence and consistency of the efficiency model employed in this study.

4.3 Phase II results

In order to implement the portfolio selection models in the second phase, the monthly returns of the 40 qualified assets resulted from the first phase are extracted from the TSE using the constructing period mentioned (a total of 520 observations). Also, the out of sample information were obtained using the testing period mentioned (a total of 280 observations). The final data regarding the return and risk factors of the qualified assets as well as the data of the other factors introduced in Sect. 3.3 are presented in Table 6.

Without loss of generality, it is assumed that $v_q = w_q = 1, \forall q$ in all the experiments (i.e. each goal is of equal importance). Also, the target values are considered to be the optimal objective function value of the individual single objective models, with each factor considered as the single objective. Moreover, the trade-off parameter in the EGP models (α) is considered to be 0.5 in all of the experiments as a fair trade-off between optimization and balance underlying philosophies of the EGP structure (Jatuphatwarodom et al., 2018). With these descriptions, the results of different robust EGP models for the multiple criteria portfolio selection problem are presented in the following sections. Also, the weights of the assets in the optimal portfolios are provided in Tables 10, 11 and 12.

Table 6 The data of factors utilized in the portfolio optimization models

Asset	Efficiency score	Industry	Factors						
			Return(RE)	Risk(RI)	Liquidity(LI)	Leverage(LE)	Activity(AC)	Profitability(PR)	
X1	1.990	Banks and credit institutions	- 0.035	0.244	0.020	0.055	0.156	0.424	
X2	1.966	Technical and engineering	0.053	0.197	0.003	0.080	0.057	- 0.013	
X3	1.533	Pharmaceutical	0.052	0.165	0.461	0.001	0.076	0.045	
X4	1.380	Investments	0.048	0.098	0.280	0.062	0.177	- 0.007	
X5	1.228	Investments	0.208	0.516	0.206	0.006	0.086	0.023	
X6	1.220	Artistic, entertainment and creative activities	0.137	0.232	0.000	0.010	0.006	0.034	
X7	1.187	Financial and monetary intermediation	0.091	0.196	0.711	0.001	0.024	0.062	
X8	1.162	Computers	0.021	0.147	0.015	0.068	0.073	0.010	
X9	1.161	Chemical	0.031	0.178	0.221	0.012	0.031	0.102	
X10	1.136	Retail	0.126	0.121	0.002	0.111	0.096	0.073	
X11	1.136	Petroleum products	- 0.004	0.213	0.002	0.057	0.065	0.039	
X12	1.124	Computers	0.123	0.170	0.012	0.147	0.187	0.061	
X13	1.111	Financial and monetary intermediation	0.215	0.505	0.055	0.003	0.019	0.035	
X14	1.107	Petroleum products	0.014	0.185	0.031	0.043	0.076	0.040	
X15	1.104	Food except sugar	0.137	0.273	0.001	0.046	0.040	0.044	
X16	1.092	Sugar	0.102	0.317	0.007	0.613	0.017	0.187	
X17	1.080	Retail	0.267	0.301	0.001	0.146	- 0.014	- 0.016	

Table 6 (continued)

Asset	Efficiency score	Industry	Factors						
			Return(RE)	Risk(RI)	Liquidity(LI)	Leverage(LE)	Activity(AC)	Profitability(PR)
X18	1.077	Base metals	0.088	0.193	0.005	0.046	0.080	0.027
X19	1.074	Automobile and parts	0.020	0.154	0.288	0.001	0.007	0.018
X20	1.073	Hotel and restaurant	0.077	0.255	0.002	0.010	0.003	0.014
X21	1.055	Non-metallic mineral	0.135	0.339	0.003	0.079	0.003	- 0.057
X22	1.044	Chemical	0.042	0.149	0.032	0.008	0.009	0.034
X23	1.032	Non-metallic mineral	0.139	0.223	0.001	0.071	0.008	0.116
X24	1.029	Financial and monetary intermediation	0.009	0.108	0.093	0.003	0.017	0.026
X25	1.025	Sugar	0.099	0.141	0.004	0.137	0.034	0.068
X26	1.025	Sugar	0.051	0.222	0.001	0.076	0.044	0.037
X27	1.025	Petroleum products	0.033	0.161	0.001	0.073	0.028	0.047
X28	1.023	Coal mining	- 0.002	0.163	0.045	0.051	0.019	0.038
X29	1.019	Sugar	0.192	0.291	0.002	0.111	0.041	0.010
X30	1.015	Food except sugar	0.208	0.270	0.001	0.104	0.036	0.017
X31	1.015	Transportation, warehousing and Communication	0.003	0.349	0.001	0.182	0.015	- 0.033
X32	1.014	Chemical	0.087	0.207	0.004	0.048	0.018	0.070

Table 6 (continued)

Asset	Efficiency score	Industry	Factors						
			Return(RE)	Risk(RI)	Liquidity(LI)	Leverage(LE)	Activity(AC)	Profitability(PR)	
X33	1.012	Petroleum products	- 0.021	0.273	0.004	0.068	0.035	0.055	
X34	1.012	Real estate	0.054	0.133	0.003	0.051	0.004	0.010	
X35	1.009	Chemical	0.105	0.134	0.030	0.034	0.030	0.074	
X36	1.009	Electrical devices	0.086	0.231	0.005	0.019	0.015	0.035	
X37	1.009	Machinery and equipment	0.098	0.159	0.004	0.111	0.093	0.052	
X38	1.009	Pharmaceutical	0.094	0.183	0.074	0.024	0.016	0.091	
X39	1.009	Pharmaceutical	0.099	0.164	0.006	0.022	0.008	0.045	
X40	1.008	Electrical devices	0.081	0.368	0.066	0.034	0.013	0.115	
		Mean	0.084	0.223	0.068	0.071	0.044	0.051	
		SD	0.069	0.095	0.145	0.099	0.047	0.074	
		Max	0.267	0.516	0.711	0.613	0.187	0.424	
		Min	- 0.035	0.098	0.000	0.001	- 0.014	- 0.057	

4.3.1 Effects of conservatism and perturbation levels in the REGP models

To have a better understanding of how the conservatism degree and data perturbation affect the GP achievement functions and the resultant portfolios, several risk levels have been established based on possible linked combinations of conservatism levels with data perturbations, which can nearly encompass optimistic, realistic and pessimistic decision making approaches. Toward this end, the perturbation range (i.e. Δ) is assumed to be 5%, 10% and 20% while the uncertainty budget is set to 0, 1, 2, 3, 4, 5, 10, 20, and 40. The conservatism level is considered to be the same for all of the uncertain factors (i.e. $\Gamma_{RE} = \Gamma_{RI} = \Gamma_{LI} = \Gamma_{LE} = \Gamma_{AC} = \Gamma_{PR} = \Gamma$) and it is assumed that the uncertain parameters vary concurrently.

According to the results presented in Tables 8 and 9 for the multi-factor REGP models, the achievement (objective) function value of the robust models is higher than that of the corresponding deterministic models due to the costs incurred for improving the model stability, which is consistent with similar applications of RGP models such as Ghahtarani and Najafi (2013), Hanks et al. (2017), and Ghasemi Bojd and Koosha (2018). This becomes more tangible when the level of conservatism or perturbation increases. For instance, a highly conservative and risk-averse version of the robust model (i.e., a conservatism level equal to 40 and a perturbation level of 20%) gives us decisions that lead to a 15.8% [= 100 * (0.578 -0.499/0.499)] higher achievement function value for REGP via combined interval and polyhedral uncertainty set and a 52.7% higher achievement function value for REGP via polyhedral uncertainty set. Lower uncertainty levels, on the other hand, increase the stability of the solutions at a more plausible achievement function value by balancing robustness and cost. Figures 6 and 7 depict how the value of the achievement function is affected by different degrees of conservatism and variations in uncertain parameters for the multi-factor REGP models via combined interval and polyhedral and polyhedral uncertainty sets, respectively. The normalized deviation $\frac{Z_R - Z_D}{Z_D}$ of the optimal value of the achievement function is used in all experiments where Z_R and Z_D are the optimal value of achievement function for the robust and deterministic models, respectively. In addition to the changes in the value of the achievement function, Fig. 6 also shows the probabilities of constraint violation for different conservatism degrees.

Figure 6 reveals that the worst-case objective function value of the REGP models via combined interval and polyhedral uncertainty set (Bertsimas and Sim (2004) approach) results when the conservatism levels of the uncertain parameters are less than their extreme values, i.e. $\Gamma = 4 < |J_i| = 40$ for $\Delta = 5\%$ and $\Delta = 10\%$, and $\Gamma = 10 < |J_i| = 40$ for $\Delta = 20\%$. Therefore, Γ does not need to be adjusted to values greater than 4 and 10 to obtain the most conservative results. On the other hand, Fig. 7 depicts that the worst-case objective function value of the REGP models via polyhedral uncertainty set is reached when the conservatism levels of the uncertain factors reach their highest values, i.e. $\Gamma = |J_i| = 40$. The initial implications are now apparent. The higher the conservatism level of uncertain parameters is in the REGP model via the polyhedral uncertainty set, the greater the impact will be on the value of the achievement function. This is not the case for variations in parameters in the REGP model via combined interval and polyhedral uncertainty set, where an increase in the level of conservatism causes less impact on the achievement function value.

According to the sensitivity analysis depicted in Figs. 6 and 7, a second insight can be drawn by comparing the extent of achievement value deterioration caused by variations in uncertain parameters. When $\Delta = 20\%$, variation in parameters has a high deterioration impact of 52.7% on the value of achievement function, for the multi-factor REGP model
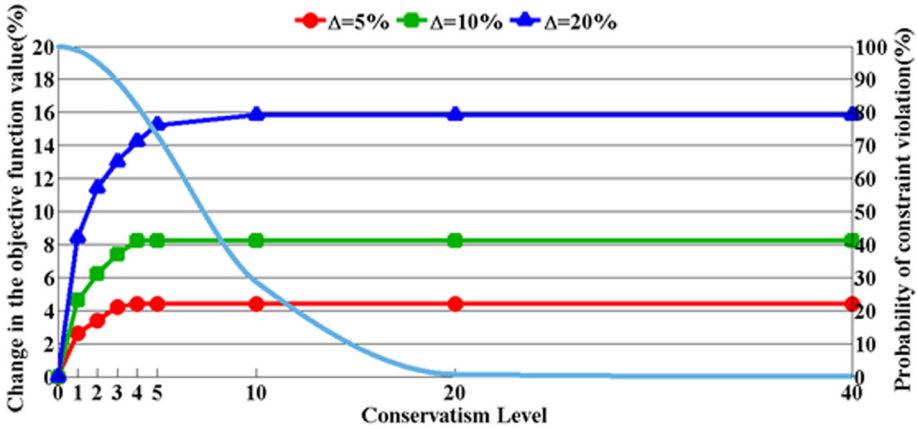
Fig. 6 Sensitivity of the achievement function to variations in factors for the multi-factor REGP model via combined interval and polyhedral uncertainty set
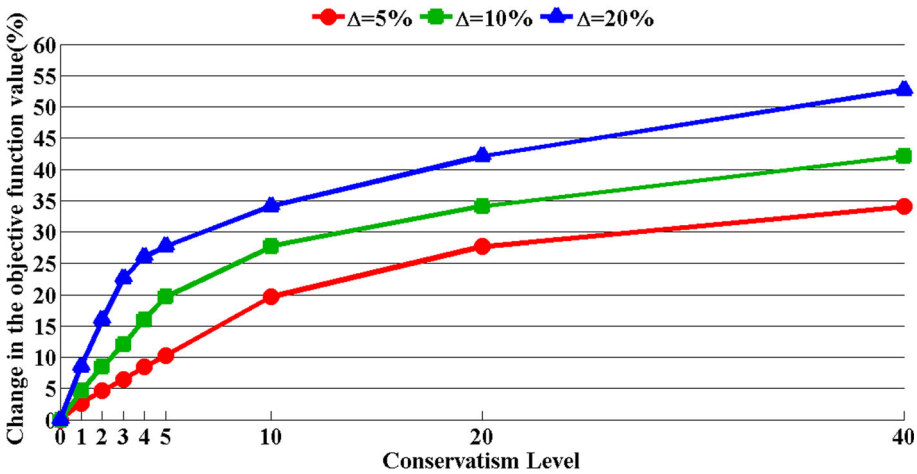

Fig. 7 Sensitivity of the achievement function to variations in factors for the multi-factor REGP model via polyhedral uncertainty set

via polyhedral uncertainty set. For the same situation of $\Delta = 20\%$ for the multi-factor REGP model via combined interval and polyhedral uncertainty set, the impact on the value of achievement function imposed by parameters variation is approximately 15.8%. Therefore, the primary focus of the FDM should be placed on more accurate estimation of input parameters data when utilizing the robust models via polyhedral uncertainty set, as variations in parameters can have the highest influence on overall achievement function value.

It should be noted that the trends of variations in the achievement functions of the 2-factor REGP models (i.e., considering only return and risk), as a sample of an REGP model with two goals, compared to their deterministic equivalents are the same as the multi-factor REGP models via both uncertainty sets. Hence, they aren't discussed here.

The results also demonstrate that despite increasing the problem size by introducing new variables and constraints, the robust approaches retained the computational tractability of the original problem as all of the robust models were solved in less than a second.
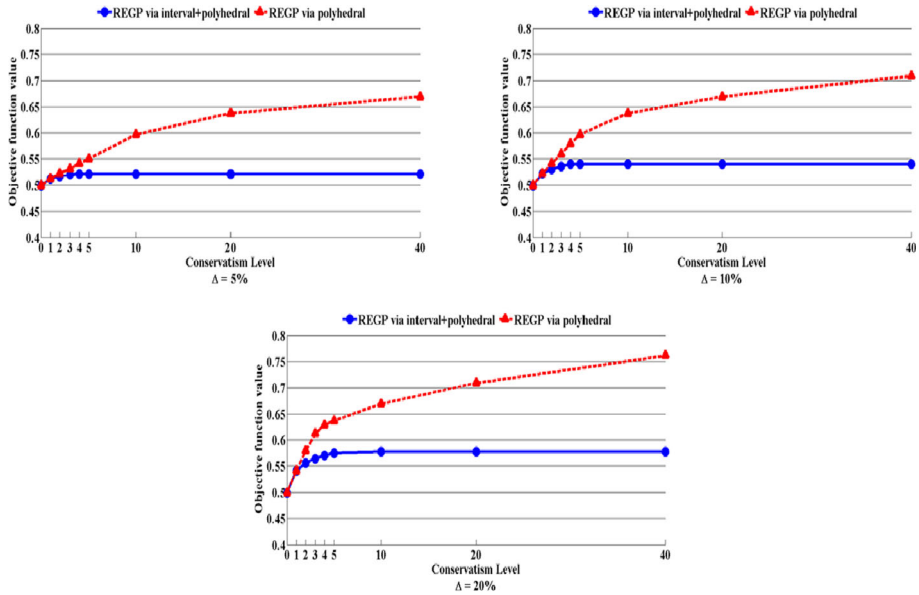
Fig. 8 Achievement function values of multi-factor REGP models via polyhedral and combined interval and polyhedral uncertainty sets for different conservatism & data perturbation levels

4.3.2 Comparison between the achievement functions of the REGP models under the two uncertainty sets

As graphically depicted in Fig. 8, under the same conservatism level, the achievement function values of the REGP model via the polyhedral uncertainty set are greater than or equal to those of the REGP model via the combined interval and polyhedral uncertainty set. The gap between the two models is relatively small for low Γ values but widens as Γ increases. In other words, the robust solutions obtained using the polyhedral uncertainty set exhibit a higher deterioration in individual objective function values and distance from the target levels compared to the REGP model using the combined interval and polyhedral uncertainty set. This suggests that the REGP model via the combined interval and polyhedral uncertainty set can offer robust solutions at a lower cost. The observation that the solution based on the polyhedral set is equal or worse than the solution based on the "interval + polyhedral" set can be attributed to the fact that combining the polyhedral set with the interval set results in a smaller uncertainty set, leading to a less conservative outcome. In terms of Δ, it can also be seen that the gap between the two models gets slightly wider as Δ increases.

4.3.3 Out-of-sample portfolio performance

To gain insight into the impacts of considering uncertainty on out-of-sample portfolio performance, the results of the deterministic EGP models with two factors (i.e., only return and risk) and with multiple factors (all the extended factors) with preferential weights and target values as previously described, as well as the results of the Markowitz model are firstly presented in Table 7. We also note that, as an ideal choice consistent with the goal target value scenarios we selected for the EGP models, we opted for the maximum return value among

Table 7 The results of the deterministic models

Performance metric	Markowitz	2-factor EGP	Multi-factor EGP
R_P	0.109	0.130	0.208
σ_P	0.223	0.172	0.212
SR	0.490	0.752	0.980
N_x	2	2	2
N_z	N.A	0	0
Obj	0.090	0.127	0.500

R_P Out of sample portfolio return, σ_P Out of sample portfolio standard deviation, SR Sharpe ratio, N_x Number of stocks selected for investment, N_z Number of objectives achieved, Obj Objective function value of the corresponding mathematical model, N.A. Not Applicable

the 40 candidate assets, which is equal to 0.267 according to Table 6, as the lower bound of the expected portfolio rate of return for the Markowitz model experiment.

The performance metrics used for assessing the results include the following: out-of-sample portfolio return, out-of-sample portfolio standard deviation, Sharpe ratio, number of stocks selected for investment, number of objectives achieved, and the objective function value of the corresponding mathematical model. Among these performance metrics, the Sharpe ratio (Sharpe, 1966), one of the most commonly used measures for evaluating out-of-sample performance, is defined as follows:

$$SR = \frac{E(R_p) - R_f}{\sigma(R_p)} \tag{65}$$

where $E(R_p)$, R_f, and $\sigma(R_p)$ denote the average out-of-sample portfolio return, risk-free return, and out-of-sample standard deviation of portfolio return, respectively.

Tables 8 and 9 present the extensive out-of-sample results of the multi-factor REGP models via combined interval and polyhedral and polyhedral uncertainty sets, respectively. These tables provide the results for each performance metric across various combinations of conservatism levels and data perturbation scenarios.

4.3.4 Discussion of the phase II results

In what follows, we discuss the results in light of the performance metrics, while also conducting a comparative analysis between nominal (deterministic counterpart) and robust models to elucidate disparities. Additionally, we undertake a meticulous comparison between the multi-factor and two-factor models, both nominal and robust, aiming to derive insights into the significance of considering multiple financial factors. According to Table 7 for the deterministic models results, and having taken into account the consistency of the number of stocks selected for investment (i.e. 2) across all three models, our observations show that the EGP models with multiple factors clearly outperformed the two-factor models. This performance disparity, notably evident against the Markowitz model, becomes most apparent in terms of out-of-sample portfolio rate of return and Sharpe ratio, to the extent that the corresponding Sharpe ratio of the multi-factor EGP model (i.e. 0.980) is twice the corresponding Sharpe ratio of the Markowitz model (i.e. 0.490).

According to Tables 8 and 9, featuring the results of the multi-factor robust models which are our main focus in this research, as expected, with $\Gamma = 0$ the results of the two REGP

Table 8 The results of the multi-factor REGP models via combined interval and polyhedral uncertainty set

Performance metric	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
R_P	5%	0.208	0.217	0.211	0.215	0.195	0.195	0.195	0.195	0.195
	10%	0.208	0.224	0.216	0.216	0.196	0.196	0.196	0.196	0.196
	20%	0.208	0.220	0.213	0.207	0.205	0.203	0.196	0.196	0.196
σ_P	5%	0.212	0.150	0.149	0.136	0.151	0.151	0.151	0.151	0.151
	10%	0.212	0.153	0.131	0.127	0.152	0.152	0.152	0.152	0.152
	20%	0.212	0.118	0.136	0.140	0.144	0.142	0.153	0.153	0.153
SR	5%	0.980	1.443	1.417	1.576	1.290	1.290	1.290	1.290	1.290
	10%	0.980	1.459	1.648	1.708	1.289	1.289	1.289	1.289	1.289
	20%	0.980	1.861	1.563	1.477	1.418	1.432	1.287	1.287	1.287
N_x	5%	2	5	5	6	2	2	2	2	2
	10%	2	5	6	6	2	2	2	2	2
	20%	2	5	7	10	11	12	2	2	2
N_z	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0
Obj	5%	0.499	0.512	0.516	0.520	0.521	0.521	0.521	0.521	0.521
	10%	0.499	0.522	0.530	0.536	0.540	0.540	0.540	0.540	0.540
	20%	0.499	0.541	0.556	0.564	0.570	0.575	0.578	0.578	0.578

R_P Out of sample portfolio return, σ_P Out of sample portfolio standard deviation, SR Sharpe ratio, N_x Number of stocks selected for investment, N_z Number of objectives achieved, Obj Objective function value of the corresponding mathematical model

Table 9 The results of the multi-factor REGP models via polyhedral uncertainty set

Performance Metric	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
R_P	5%	0.208	0.217	0.224	0.219	0.220	0.225	0.206	0.168	0.171
	10%	0.208	0.224	0.220	0.226	0.228	0.206	0.168	0.171	0.186
	20%	0.208	0.220	0.228	0.203	0.176	0.168	0.171	0.186	0.186
σ_P	5%	0.212	0.150	0.153	0.116	0.118	0.125	0.113	0.050	0.062
	10%	0.212	0.153	0.118	0.126	0.129	0.113	0.050	0.062	0.072
	20%	0.212	0.118	0.129	0.111	0.071	0.050	0.062	0.072	0.082
SR	5%	0.980	1.443	1.459	1.885	1.861	1.799	1.818	3.334	2.751
	10%	0.980	1.459	1.861	1.789	1.763	1.818	3.334	2.751	2.571
	20%	0.980	1.861	1.763	1.834	2.473	3.334	2.751	2.571	2.262
N_x	5%	2	5	5	5	5	4	5	13	17
	10%	2	5	5	4	4	5	13	17	26
	20%	2	5	4	5	10	13	17	26	33
N_z	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0
Obj	5%	0.499	0.512	0.522	0.531	0.541	0.550	0.597	0.637	0.669
	10%	0.499	0.522	0.541	0.559	0.579	0.597	0.637	0.669	0.709
	20%	0.499	0.541	0.579	0.612	0.629	0.637	0.669	0.709	0.762

R_P Out of sample portfolio return, σ_P Out of sample portfolio standard deviation, SR Sharpe ratio, N_x Number of stocks selected for investment, N_z Number of objectives achieved, Obj Objective function value of the corresponding mathematical model

models are the same and equal to the nominal model. This could be seen as a sign of robust model validation. Notably, a Γ of 0 indicates absence of uncertainty consideration, thereby rendering the model susceptible to variations in input parameters. For $\Gamma = 1$, it is observed that the "polyhedral" and the "combined interval and polyhedral" set-based solutions are also identical in terms of all performance metrics, as observed earlier in subSect. 4.3.2 regarding achievement function values. This is because the corresponding uncertainty sets are also identical when $\Gamma \leq 1$, as shown in Fig. 4 . As $\Gamma \geq 1$, the solutions based on the combined uncertainty set outperform those based on the pure polyhedral set in terms of the achievement function. This superiority arises from the smaller uncertainty sets associated with the combined set, because of the imposition of the bounded box constraint, in contrast to the pure polyhedral set, where the corresponding achievement function values deteriorate rapidly. In terms of portfolio returns, both models generally lead to portfolios having relatively high returns across all uncertainty settings, surpassing the portfolio return obtained from the Markowitz model (with $R_P = 0.109$). As far as the out-of sample risk values are concerned, both of the REGP models consistently yield portfolios with lower risk compared to the Markowitz model (with $\sigma_P = 0.223$). Comparing the REGP models via both uncertainty sets against their nominal counterparts (for which the results are presented in the last column of Table 7 or in Tables 8 and 9 with $\Gamma = 0$) in terms of out-of-sample risk and return performance unveils intriguing dynamics. While both higher and lower returns are realized by robust models (compared to their nominal counterpart model with $R_P = 0.208$), risk values consistently favor robust models (compared to their nominal counterpart model with $\sigma_P = 0.212$) across all uncertainty levels. Also, under varying degrees of conservatism, we observe a reduction in returns for both of the REGP models relative to their nominal counterpart, particularly pronounced in the highly conservative versions of the models, which is more notable for the REGP model via polyhedral uncertainty set. In this regard, it is also observed that the REGP model utilizing the polyhedral uncertainty set consistently produces portfolios with significantly lower risk values compared to the nominal model, particularly with its highly conservative and risk-averse versions (marked by higher linked levels of Γ and Δ). This could be attributed to the issue that as we increase the level of conservatism in robust optimization models, they become more risk-averse, aiming to protect against worst-case scenarios. This increased aversion to risk may lead to the selection of assets with lower expected returns but lower associated risk, resulting in reduced overall portfolio returns. Consequently, the observed changes in out-of-sample risk and return values for REGP models via both uncertainty sets result in portfolio Sharpe ratios that consistently surpass those of the Markowitz model ($SR = 0.490$) and the nominal EGP model ($SR = 0.980$) across all uncertainty levels.

Regarding the influence of perturbation levels (Δ) of the uncertain parameters on the aforementioned performance metrics, it is observable that while changes in return, risk, and Sharpe ratio values remain relatively modest for the REGP model using the combined interval and polyhedral uncertainty set at constant conservatism degrees, the influence of Δ value variations is more pronounced for the REGP model employing the polyhedral uncertainty set. From the above observations, it can be concluded that the sensitivity of the REGP model via the pure polyhedral uncertainty set is higher in terms of changes in performance metric values against changes in uncertainty levels. This heightened sensitivity was shown in sub-Sects. 4.3.1 and 4.3.2 concerning the achievement function values, and it can be seen from the results that it also extends to other metrics such as N_x.

Concerning the number of selected stocks (N_x), two general viewpoints prevail. A smaller number of selected stocks will result in lower transaction costs and easier portfolio management. On the other hand, a larger number of stocks also brings with it the benefits of

diversification (Georgantas et al., 2021). It is evident from the results that the deterministic models, i.e. the Markowitz and the nominal EGP models, along with some conservative versions of the REGP model employing the combined interval and polyhedral uncertainty set lead to the sparsest portfolios, each consisting of 2 assets. In stark contrast, The REGP model employing the pure polyhedral uncertainty set results in the largest and most diversified portfolios, comprising 33 assets in the highly conservative and risk-averse version of the model. Moreover, in terms of portfolio diversification, we observe that the difference between the nominal and robust models is more trivial in the REGP model via combined interval and polyhedral uncertainty set, with a few exceptions for certain levels of conservatism degree. This phenomenon is not unusual since the application of RO does not mean increased diversification compared to the nominal model. Instead, it tries to create immunization against the worst-case realizations of the uncertain input parameters. In this respect, if a trade-off is made with high return potentials with low risk potentials in nominal models, it will happen with the worst-case return and risk potentials in the robust counterpart problem (Xidonas et al., 2020). Therefore, if the nominal model generates portfolios that are not initially diversified, then the robust model might not necessarily improve diversification and, under certain conditions involving undesirable large perturbations in uncertain parameters like return values, might even lead to less desirable diversification. However, it is seen from the results that the REGP model via the polyhedral uncertainty set offers diversification benefits when compared to the nominal model and also to the REGP models via combined interval and polyhedral uncertainty set, which is especially pronounced in the highly conservative and risk-averse version of this robust model.

Additionally, regarding the number of goals achieved, both models fell short of meeting all the target values. This outcome can be explained through the additional costs incurred in stabilizing the model and making it robust against infeasibility. As the achievement function of the GP models comprises the summation of deviations from targets and the maximum deviation, an increase in this value indicates either a deterioration of objective functions or a deviation of previously achieved goals from their respective targets within the robust model. As the number of goals achieved in the deterministic models is also zero, the first case has occurred—wider gaps between objective functions and their designated targets in the robust models. It should be noted that N_z is an important parameter in the output of GP models as it may account for the Pareto efficiency or inefficiency status of the resulting solution. If all targets were met (E.g. $N_z = 6$ in the case of multi-factor models), this would most likely indicate a Pareto inefficient solution, and a Pareto efficiency restoration technique may need to be employed in order to restore Pareto efficiency to the solution (Jones & Tamiz, 2010). Therefore, with $N_z = 0$, there is less concern about possible Pareto inefficiency within the solutions. It can hence give us the assurance that we are dealing with well-constructed GP models for both deterministic and uncertain conditions.

Altogether, it can be deduced from Table 7, 8, and 9 that both multi-factor REGP models consistently outperformed the well-established Markowitz model across all of the uncertainty levels. Furthermore, these multi-factor REGP models demonstrated superiority over their deterministic counterpart, the multi-factor EGP model, in many of the uncertainty levels, underscoring their efficacy in navigating financial uncertainties. It's worth noting that the REGP model via polyhedral uncertainty set achieved this aim while incurring a higher degree of deviation from the target values compared to its counterpart that utilized the combined interval and polyhedral uncertainty set (Bertsimas and Sim (2004) approach).

To derive insight into the impact of multiple factors (the extended factors) considered on portfolio performance, we conducted an extensive comparison of out-of-sample performance

metrics in terms of returns and risks between the multi-factor REGP models via both uncertainty sets, and their corresponding 2-factor models, which considered only return and risk factors, as well as the classic Markowitz model. Figures 9, 10, 11 and 12 provide a visual representation of these comparisons, with a reminder that the bars with $\Gamma = 0$ indicate the deterministic equivalent of each robust model in these assessments.

As evidenced by the charted results, the findings clearly demonstrate the superiority of the deterministic equivalent models, and both the multi-factor and 2-factor REGP models via both uncertainty sets over the traditional Markowitz model in terms of out-of-sample portfolio returns and standard deviations across all designed uncertainty levels. This substantiates the effectiveness of the proposed GP approaches, encompassing both multi-factor
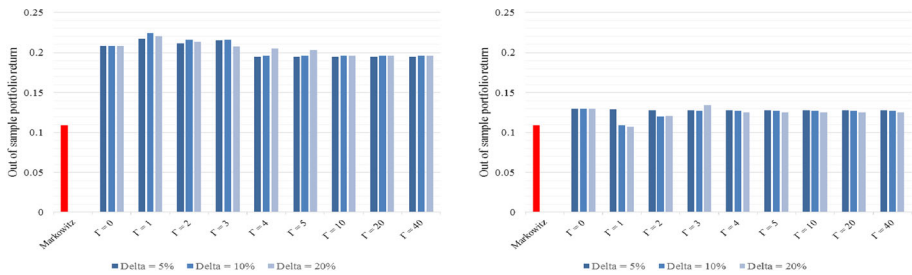

Fig. 9 Out-of-sample portfolio returns for multi-factor (left) and 2-factor (right) REGP models via interval + polyhedral uncertainty set
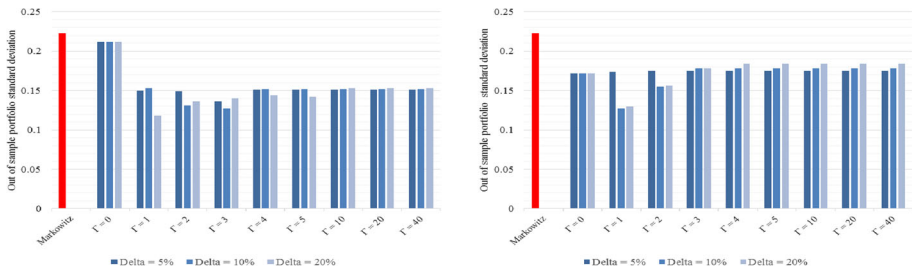

Fig. 10 Out-of-sample portfolio standard deviations for multi-factor (left) and 2-factor (right) REGP models via interval + polyhedral uncertainty set
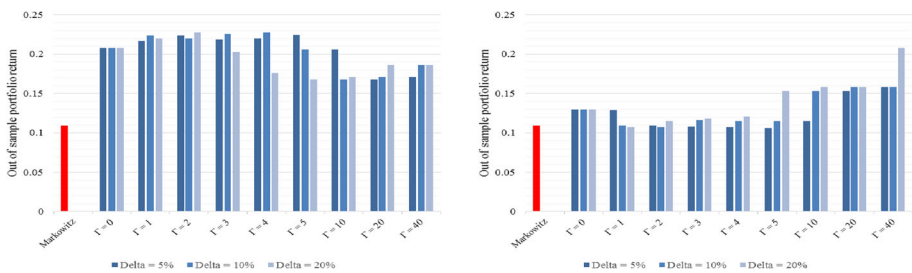

Fig. 11 Out-of-sample portfolio returns for multi-factor (left) and 2-factor (right) REGP models via polyhedral uncertainty set
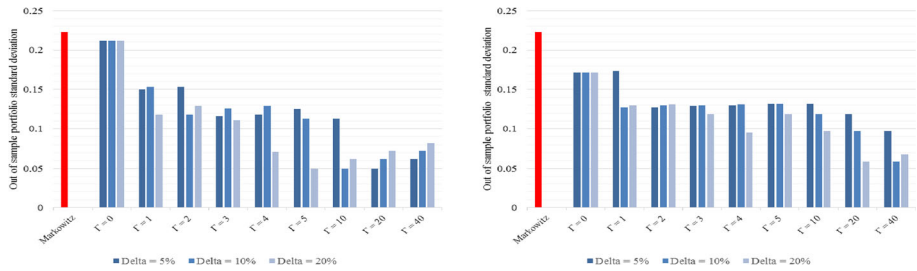
Fig. 12 Out-of-sample portfolio standard deviations for multi-factor (left) and 2-factor (right) REGP models via polyhedral uncertainty set

and two-factor nominal and robust models, in addressing the complexities of stock portfolio selection compared to the classic Markowitz model. Moreover, it is evident that the multi-factor REGP models, irrespective of the uncertainty set used, delivered higher returns and exhibited relatively lower standard deviations when compared to their 2-factor counterparts across various combinations of uncertainty levels. The outperformance of the multi-factor models, which in this study incorporate a broader range of fundamental financial factors in addition to traditional risk and return factors, over their 2-factor counterparts could be attributed to several reasons. The inclusion of fundamental financial factors adds an additional layer of diversification potential to the portfolio. These factors can capture nuances in the financial health and stability of assets that may not be fully reflected in simple risk and return metrics. This diversification could help mitigate risks associated with individual assets and enhance the overall stability of the portfolio. Moreover, fundamental financial factors often provide insights into the intrinsic value and financial stability of assets. During periods of market volatility or economic downturns, assets with strong fundamental characteristics tend to be more resilient. By including these factors, the multi-factor models are better equipped for turbulent market conditions, leading to more stable returns and lower standard deviations. Further, fundamental financial factors are often associated with the long-term financial health and performance of assets. By taking a longer-term perspective, the multi-factor models can identify assets that are not only attractive in the short term but also have the potential for sustained performance over time. This long-term view could contribute to more consistent returns and lower standard deviations. Also, regarding the comparison of the REGP models against their nominal counterparts, as discussed earlier, out-of-sample risk values consistently favored robust models for the multi-factor REGP models. The extent of this superiority is not observed in the robust 2-factor models compared to their deterministic equivalents.

Overall, the results mentioned above align closely with the findings of Tamiz and Azmi (2019), affirming the advantages of incorporating additional financial factors rooted in fundamental analysis alongside conventional technical risk and return factors within the EGP framework for optimizing stock portfolios. Importantly, these benefits are evident not only in deterministic settings but also when confronted with the inherent uncertainties of financial markets, a facet notably lacking in Tamiz et al. (2013) and Tamiz and Azmi (2019). From the perspective of applying the GP methodology to portfolio selection, Tamiz et al. (2013) introduced a practical and effective approach, allowing decision-makers to incorporate their preferred factors, termed extended factors, and ideal aspiration levels into their choice of GP model to construct their desired portfolios, particularly in the context of international mutual funds. This framework was subsequently adapted for stock portfolio selection in Tamiz and

Azmi (2019), yielding promising outcomes. In our study, we have taken this framework further, applying it to a different market dataset, incorporating distinct factors, and employing a unique GP variant, the EGP. Importantly, we have extended its application to encompass scenarios where financial market uncertainties are significant, making our framework more realistic for real-world use compared to prior works. Furthermore, the choice of the EGP variant, with its parametric analysis capabilities for trade-offs between efficiency and balance, offers flexibility and advantages over GP variants used in Tamiz et al. (2013) and Tamiz and Azmi (2019). This choice of GP variant is pivotal, as various GP variants yield different trade-offs between objectives, target values, and the penalization of unwanted deviational variables, which may lead to the construction of better and more diverse portfolios. The positive findings presented here underscore the framework's practicality and suggest its potential applicability to a range of other financial instruments, such as bonds, ETFs, and more. Also, extending the framework to incorporate various factors, which could characterize today's world and favorably influence investors' investment preferences, holds promise.

RGP is a relatively underdeveloped, less widely researched, and less applied topic, especially when considering the various robust concepts and variants of the GP technique (Hanks et al., 2017; Mensah & Rocca, 2019). Our study contributes to the growing literature on RGP by developing and analyzing the REGP framework, highlighting its potential as a versatile tool that offers a pragmatic, flexible, and straightforward approach to tackling complex multi-objective optimization problems in uncertain environments. EGP underpins REGP, enabling a parametric analysis of trade-offs between efficiency and balance among target values, which makes REGP highly flexible. Our proposed REGP framework in this study features two uncertainty sets—polyhedral and combined interval and polyhedral, which yields robust optimal decisions while preserving linearity and computational tractability, distinguishing it from previous works. It adapts solutions across a spectrum from worst-case to deterministic outcomes based on adjustable parameters associated with the uncertainty sets. The findings illustrated that total deviations from decision makers' set targets are dependent on the type of the uncertainty set used, unlike the findings of Mensah and Rocca (2019). Furthermore, REGP excels by accommodating uncertainties not only in goal function coefficients and target values but also in system constraints—a critical feature for achieving complete robustness in problems where both constraint parameters and goals are imprecisely defined. Additionally, REGP can be seamlessly extended and combined with other uncertainty modeling techniques, depending on the nature of the uncertainty source, all within a single coherent framework. In this study, REGP was applied to portfolio selection, a context well-suited due to its inherent multi-criteria nature and the presence of uncertainties. REGP can address both dimensions effectively, offering advantages over the classic MV framework, which cannot handle uncertainties or diverse investor preferences. However, the REGP framework has potential applications in other contexts as a versatile tool for robust multi-objective decision making, which is strongly recommended.

5 Conclusions

This paper has proposed an innovative framework tailored for constructing asset portfolios under conditions of data uncertainty. Comprising two distinct phases, the proposed framework harnesses the power of two prominent decision analytical methods of DEA and GP to achieve its objective. The initial phase employs a super-efficiency DEA model to streamline asset selection by identifying suitable investment candidates based on the efficiency characteristic of the assets. An extensive set of financial parameters was utilized as inputs

and outputs of the DEA model for efficiency evaluation. This phase demonstrates its merit through a dual advantage: firstly, it mitigates the computational complexity of the Portfolio selection problem through eliminating the need to deal with an extensive pool of assets. Secondly, a thorough assessment of assets has been carried out by the end of this phase and investments with promising potential are selected. Subsequently, the second phase leverages a multi-objective mathematical model to determine the optimal asset weights within the portfolio, with each objective tailored to reflect specific financial factors spanning both technical and fundamental aspects. The multi-objective problem was then handled using the EGP methodology, affording investors the flexibility to accommodate their nuanced preferences pertaining to both financial considerations and the underlying philosophies of weighted or MinMax GP variants. Furthermore, acknowledging the intrinsic volatility of financial market parameters, the study pioneers the development of robust counterparts for the EGP models. These robust models, developed via a polyhedral and a combined interval and polyhedral uncertainty sets, offer optimal decisions which are robust against parameter inaccuracies and slight fluctuations. Remarkably, this robustness is achieved while preserving linearity and tractability. The proposed framework enables investors to encompass their unique preferences and conservatism levels, whilst optimizing portfolios amid data uncertainty.

5.1 Implications

The applicability and efficacy of the proposed approach have been illustrated through a comprehensive case study of the Tehran stock exchange. The empirical investigation yielded several implications. Firstly, the implementation of the super-efficiency DEA approach in the initial phase displayed the capacity to efficiently identify optimal investment candidates. This capability not only assists investors and portfolio managers in reducing the computational complexities and distractions in stock markets where a large number of assets are registered, but also offers insights into the underlying business strength of the firms, thus facilitating deeper analysis. Furthermore, the integration of multiple factors (called extended factors), encompassing both fundamental analysis-based metrics and traditional technical risk-return factors, within the portfolio optimization framework led to the formulation of portfolios that exhibited superior out of sample performance compared to conventional and two-factor strategies. This performance superiority extended across deterministic scenarios and situations marked by uncertainty, underscoring the strength and resilience of the proposed approach. Additionally, the flexibility of the framework was underscored by its ability to incorporate a diverse array of factors, ranging from financial to non-financial, alongside customizable preference weights in the corresponding GP model's achievement function, goal target values, and conservatism levels. This adaptability ensures the alignment of the approach with the specific preferences and utilities of investors in curating their optimal asset portfolios. Also, the EGP variant along with its robust counterparts proved to be a promising tool in portfolio selection decisions under both deterministic and uncertain conditions. The robust EGP portfolio optimization models generally outperformed their nominal counterpart and the deterministic Markowitz models, highlighting their robustness and dependability under varying market circumstances. Moreover, a detailed comparison between the REGP model via polyhedral and combined interval and polyhedral uncertainty sets revealed intriguing insights. While the former displayed higher total deviations from set targets, it concurrently offered diversification benefits and more favorable out-of-sample risk values. This revelation underscores the importance of well-informed model selection according to the investors' preferences. The findings illustrated that in the context of the REGP with uncertainty sets,

total deviations from decision makers' set targets are dependent on the type of the uncertainty set used. The observation regarding the disparity in total deviations from target values within the REGP models employing the two distinct uncertainty sets could offer a generalizable guideline applicable to various scenarios across different application domains where linear robust counterparts are to be employed for given linear EGP problems, involving the mentioned uncertainty sets. The analysis also leads to the conclusion that for bounded uncertainty, it is advisable to combine the uncertainty set with the interval set to prevent overly conservative solutions. Conversely, all the models possess the adaptability to fine-tune the solution along a spectrum from the worst-case scenario to the deterministic solution. This adaptability relies on the choice of adjustable parameters associated with their respective uncertainty set. Overall, the findings of the study collectively accentuate the approach's utility in addressing the intricacies of portfolio selection. The integration of DEA and GP, fortified by extended financial factors and robust optimization techniques, surpasses conventional methods. It provides a comprehensive toolkit for informed investment decision-making, tailored to address the multifaceted considerations and uncertainties inherent in contemporary financial markets.

5.2 Future research directions

For future studies, there are several directions that researchers and practitioners can consider. The EGP framework offers opportunities for investigating various priorities for the goals, as well as exploring the inclusion of other meta-objectives, like the number of goals achieved (Jones & Jimenez, 2013). Additionally, other GP variants, like meta-GP (Uría et al., 2002) and multi-choice GP (Chang, 2007), could be examined, accompanied by the development of their robust counterpart models via appropriate uncertainty sets. The versatility of REGP as a robust multi-objective decision-making methodology makes it highly promising for addressing data-driven problems in diverse contexts. Moreover, the combination of DEA and GP in the portfolio selection context can extend beyond the current study's scope, with potential for innovative integration methods. The scope can also be extended to the incorporation of robust DEA models, allowing the formulation of robust combined methodologies (Peykani et al., 2020). Another avenue for investigation is the expansion of GP models to include various factors like market ratios (e.g., EPS, P/E, dividends) for effective portfolio optimization. Furthermore, the proposed approach's applicability could be tested using data from different financial markets to assess its broader utility. Also, the interface between MCDM and Artificial intelligence/Machine Learning techniques, such as clustering, can be scrutinized to explore how it can enhance portfolio diversity while considering market uncertainties. Lastly, another avenue to explore involves evaluating the performance of portfolio optimization models under conditions involving cardinality constraints, transaction costs and multi-period time horizons.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Appendix

Weight of the Selected Assets from Phase 1 in the Optimal Portfolios:

See (Tables 10, 11 and 12).

Table 10 Weight of the assets in the optimal portfolios resulted from the deterministic models

Asset	Markowitz	2-factor EGP	Multi-factor EGP
X1	0	0	0.499
X10	0	0.409	0
X17	0.994	0.591	0.501
X30	0.006	0	0

Table 11 Weight of the assets in the optimal portfolios resulted from the multi-factor REGP models via combined interval and polyhedral uncertainty set

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X1	5%	0.499	0.354	0.358	0.328	0.248	0.248	0.248	0.248	0.248
	10%	0.499	0.356	0.330	0.328	0.250	0.250	0.250	0.250	0.250
	20%	0.499	0.311	0.322	0.343	0.346	0.336	0.255	0.255	0.255
X6	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0.073	0.077	0.069	0	0	0
X10	5%	0	0.229	0.245	0.229	0.752	0.752	0.752	0.752	0.752
	10%	0	0.208	0.197	0.167	0.750	0.750	0.750	0.750	0.750
	20%	0	0.195	0.160	0.089	0.084	0.119	0.745	0.745	0.745
X12	5%	0	0.172	0.134	0.163	0	0	0	0	0
	10%	0	0.212	0.164	0.170	0	0	0	0	0
	20%	0	0.199	0.129	0.091	0.086	0.077	0	0	0
X13	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0.030	0.029	0	0	0
X15	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0.062	0.056	0.053	0	0	0
X16	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0
X17	5%	0.501	0.107	0.115	0.092	0	0	0	0	0
	10%	0.501	0.098	0.080	0.078	0	0	0	0	0
	20%	0.501	0	0.060	0.042	0.040	0.036	0	0	0

Table 11 (continued)

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X23	5%	0	0	0	0.086	0	0	0	0	0
	10%	0	0	0.125	0.151	0	0	0	0	0
	20%	0	0.177	0.101	0.081	0.076	0.068	0	0	0
X29	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0.058	0.052	0.050	0	0	0
X30	5%	0	0.138	0.148	0.103	0	0	0	0	0
	10%	0	0.126	0.103	0.107	0	0	0	0	0
	20%	0	0.118	0.077	0.054	0.051	0.050	0	0	0
X35	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0.152	0.107	0.101	0.090	0	0	0
X37	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0.023	0	0	0

Table 12 Weight of the assets in the optimal portfolios resulted from the multi-factor REGP models via polyhedral uncertainty set

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X1	5%	0.499	0.354	0.356	0.302	0.311	0.266	0.189	0.023	0.015
	10%	0.499	0.356	0.311	0.273	0.291	0.189	0.023	0.015	0.008
	20%	0.499	0.311	0.291	0.179	0.030	0.023	0.015	0.008	0.005
X2	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.043
X5	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.009
X6	5%	0	0	0	0	0	0	0	0.080	0.066
	10%	0	0	0	0	0	0	0.080	0.066	0.036
	20%	0	0	0	0	0.106	0.080	0.066	0.036	0.028
X8	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.052
	20%	0	0	0	0	0	0	0	0.052	0.047

Table 12 (continued)

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X9	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.008
X10	5%	0	0.229	0.208	0.198	0.195	0.276	0.210	0.093	0.073
	10%	0	0.208	0.195	0.273	0.267	0.210	0.093	0.073	0.040
	20%	0	0.195	0.267	0.217	0.122	0.093	0.073	0.040	0.031
X11	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.039
X12	5%	0	0.172	0.212	0.202	0.199	0.208	0.158	0.047	0.037
	10%	0	0.212	0.199	0.206	0.201	0.158	0.047	0.037	0.020
	20%	0	0.199	0.201	0.150	0.063	0.047	0.037	0.020	0.018
X13	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.019
	20%	0	0	0	0	0	0	0	0.019	0.017
X14	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.050
	20%	0	0	0	0	0	0	0	0.050	0.045
X15	5%	0	0	0	0	0	0	0	0.068	0.056
	10%	0	0	0	0	0	0	0.068	0.056	0.036
	20%	0	0	0	0	0.090	0.068	0.056	0.036	0.028
X17	5%	0.501	0.107	0.098	0	0	0	0	0	0
	10%	0.501	0.098	0	0	0	0	0	0	0.019
	20%	0.501	0	0	0	0	0	0	0.019	0.015
X18	5%	0	0	0	0	0	0	0	0.097	0.080
	10%	0	0	0	0	0	0	0.097	0.080	0.047
	20%	0	0	0	0	0	0.097	0.080	0.047	0.043
X20	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.038
	20%	0	0	0	0	0	0	0	0.038	0.033
X22	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.066
	20%	0	0	0	0	0	0	0	0.066	0.056
X23	5%	0	0	0	0.179	0.177	0.250	0.190	0.084	0.056
	10%	0	0	0.177	0.247	0.242	0.190	0.084	0.056	0.029
	20%	0	0.177	0.242	0.196	0.111	0.084	0.056	0.029	0.020
X24	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.019

Table 12 (continued)

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X25	5%	0	0	0	0	0	0	0	0	0.058
	10%	0	0	0	0	0	0	0	0.058	0.032
	20%	0	0	0	0	0	0	0.058	0.032	0.025
X26	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.044
	20%	0	0	0	0	0	0	0	0.044	0.038
X27	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.061
	20%	0	0	0	0	0	0	0	0.061	0.047
X28	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.040
X29	5%	0	0	0	0	0	0	0	0	0.048
	10%	0	0	0	0	0	0	0	0.048	0.026
	20%	0	0	0	0	0	0	0.048	0.026	0.020
X30	5%	0	0.138	0.126	0.119	0.118	0	0	0.056	0.044
	10%	0	0.126	0.118	0	0	0	0.056	0.044	0.024
	20%	0	0.118	0	0	0	0.056	0.044	0.024	0.019
X32	5%	0	0	0	0	0	0	0	0.090	0.075
	10%	0	0	0	0	0	0	0.090	0.075	0.047
	20%	0	0	0	0	0	0.090	0.075	0.047	0.033
X33	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0
	20%	0	0	0	0	0	0	0	0	0.031
X34	5%	0	0	0	0	0	0	0	0	0
	10%	0	0	0	0	0	0	0	0	0.074
	20%	0	0	0	0	0	0	0	0.074	0.063
X35	5%	0	0	0	0	0	0	0.252	0.111	0.087
	10%	0	0	0	0	0	0.252	0.111	0.087	0.044
	20%	0	0	0	0.259	0.146	0.111	0.087	0.044	0.030
X36	5%	0	0	0	0	0	0	0	0	0.067
	10%	0	0	0	0	0	0	0	0.067	0.042
	20%	0	1	0	0	0	0	0.067	0.042	0.036
X37	5%	0	0	0	0	0	0	0	0.093	0.073
	10%	0	0	0	0	0	0	0.093	0.073	0.040
	20%	0	0	0	0	0.122	0.093	0.073	0.040	0.031
X38	5%	0	0	0	0	0	0	0	0.044	0.035
	10%	0	0	0	0	0	0	0.044	0.035	0.028
	20%	0	0	0	0	0.058	0.044	0.035	0.028	0.024

Table 12 (continued)

Asset	Perturbation level (Δ)	Conservatism degree (Γ)								
		0	1	2	3	4	5	10	20	40
X39	5%	0	0	0	0	0	0	0	0.114	0.092
	10%	0	0	0	0	0	0	0.114	0.092	0.051
	20%	0	0	0	0	0.150	0.114	0.092	0.051	0.039
X40	5%	0	0	0	0	0	0	0	0	0.039
	10%	0	0	0	0	0	0	0	0.039	0.027
	20%	0	0	0	0	0	0	0.039	0.027	0.020

References

Abad, C., Thore, S. A., & Laffarga, J. (2004). Fundamental analysis of stocks by two-stage DEA. *Managerial and Decision Economics, 25*(5), 231–241.

Almeida-Filho, A. T. D., de Lima Silva, D. F., & Ferreira, L. (2021). Financial modelling with multiple criteria decision making: A systematic literature review. *Journal of the Operational Research Society, 72*(10), 2161–2179.

Amin, G. R., & Hajjami, M. (2021). Improving DEA cross-efficiency optimization in portfolio selection. *Expert Systems with Applications, 168*, 114280.

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science, 39*(10), 1261–1264.

Andriosopoulos, D., Doumpos, M., Pardalos, P. M., & Zopounidis, C. (2019). Computational approaches and data analytics in financial services: A literature review. *Journal of the Operational Research Society, 70*(10), 1581–1599.

Aouni, B., Colapinto, C., & La Torre, D. (2014). Financial portfolio management through the goal programming model: Current state-of-the-art. *European Journal of Operational Research, 234*(2), 536–545.

Aouni, B., Doumpos, M., Pérez-Gladish, B., & Steuer, R. E. (2018). On the increasing importance of multiple criteria decision aid methods for portfolio selection. *Journal of the Operational Research Society, 69*(10), 1525–1542.

Avkiran, N. K. (2011). Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks. *Omega, 39*(3), 323–334.

Ballestero, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., & Pla-Santamaria, D. (2012). Socially responsible investment: A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research, 216*(2), 487–494.

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research, 23*(4), 769–805.

Ben-Tal, A., & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming, 88*(3), 411–424.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Princeton University Press.

Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming, 98*(1), 49–71.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research, 52*(1), 35–53.

Bettman, J. L., Sault, S. J., & Schultz, E. L. (2009). Fundamental and technical analysis: Substitutes or complements? *Accounting & Finance, 49*(1), 21–36.

Bilbao-Terol, A., Arenas-Parra, M., & Cañal-Fernández, V. (2012). Selection of socially responsible portfolios using goal programming and fuzzy technology. *Information Sciences, 189*, 110–125.

Boubaker, S., Le, T. D., Manita, R., & Ngo, T. (2023). The trade-off frontier for ESG and Sharpe ratio: a bootstrapped double-frontier data envelopment analysis. *Annals of Operations Research*, pp. 1–25.

Bravo, M., Pla-Santamaria, D., & Garcia-Bernabeu, A. (2010). Portfolio selection from multiple benchmarks: A goal programming approach to an actual case. *Journal of Multi-Criteria Decision Analysis, 17*(5–6), 155–166.

Bravo, M., Jones, D., Pla-Santamaria, D., & Salas-Molina, F. (2022). Encompassing statistically unquantifiable randomness in goal programming: an application to portfolio selection. *Operational Research*, pp. 1–22.

Chang, C. T. (2007). Multi-choice goal programming. *Omega, 35*(4), 389–396.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444.

Chen, H. H. (2008). Stock selection using data envelopment analysis. *Industrial Management and Data Systems, 108*(9), 1255–1268.

Cheng, J., Feng, X., & Bai, X. (2021). Modeling equitable and effective distribution problem in humanitarian relief logistics by robust goal programming. *Computers & Industrial Engineering, 155*, 107183.

Choobineh, M., & Mohagheghi, S. (2016). A multi-objective optimization framework for energy and asset management in an industrial Microgrid. *Journal of Cleaner Production, 139*, 1326–1338.

Colapinto, C., La Torre, D., & Aouni, B. (2019). Goal programming for financial portfolio management: A state-of-the-art review. *Operational Research, 19*(3), 717–736.

Contreras, I., Hidalgo, J. I., & Núñez-Letamendia, L. (2012, April). A GA combining technical and fundamental analysis for trading the stock market. In *European Conference on the Applications of Evolutionary Computation* (pp. 174–183). Springer, Berlin, Heidelberg.

Cook, W. D., Tone, K., & Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega, 44*, 1–4.

De, M., Mangaraj, B. K., & Das, K. B. (2018). A fuzzy goal programming model in portfolio selection under competitive-cum-compensatory decision strategies. *Applied Soft Computing, 73*, 635–646.

Deng, X., & Yuan, Y. (2021). A novel fuzzy dominant goal programming for portfolio selection with systematic risk and non-systematic risk. *Soft Computing, 25*(23), 14809–14828.

Doumpos, M., & Zopounidis, C. (2014). *Multicriteria analysis in finance*. Springer.

Dutta, P., Jain, A., & Gupta, A. (2020). Performance analysis of non-banking finance companies using two-stage data envelopment analysis. *Annals of Operations Research, 295*(1), 91–116.

Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics, 84*(2), 523–538.

Edirisinghe, N. C., & Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Journal of Banking and Finance, 31*(11), 3311–3335.

Edirisinghe, N. C. P., & Zhang, X. (2008). Portfolio selection under DEA-based relative financial strength indicators: Case of US industries. *Journal of the Operational Research Society, 59*(6), 842–856.

Emrouznejad, A., & Yang, G. L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences, 61*, 4–8.

Georgantas, A., Doumpos, M., & Zopounidis, C. (2021). Robust optimization approaches for portfolio selection: a comparative analysis. *Annals of Operations Research*, pp. 1–17.

Ghahtarani, A., Saif, A., & Ghasemi, A. (2022). Robust portfolio selection problems: a comprehensive review. *Operational Research*, pp. 1–62.

Ghahtarani, A., & Najafi, A. A. (2013). Robust goal programming for multi-objective portfolio selection problem. *Economic Modelling, 33*, 588–592.

Ghasemi Bojd, F., & Koosha, H. (2018). A robust goal programming model for the capital budgeting problem. *Journal of the Operational Research Society, 69*(7), 1105–1113.

Habib, A. M., & Kayani, U. N. (2023, June). Evaluating the Super-Efficiency of Working Capital Management Using Data Envelopment Analysis: Does COVID-19 Matter?. In *Operations Research Forum* (Vol. 4, No. 2, pp. 1–20). Springer International Publishing.

Hanks, R. W., Weir, J. D., & Lunday, B. J. (2017). Robust goal programming using different robustness echelons via norm-based and ellipsoidal uncertainty sets. *European Journal of Operational Research, 262*(2), 636–646.

Hanks, R. W., Lunday, B. J., & Weir, J. D. (2020). Robust goal programming for multi-objective optimization of data-driven problems: A use case for the United States transportation command's liner rate setting problem. *Omega, 90*, 101983.

Hendalianpour, A., Fakhrabadi, M., Zhang, X., Feylizadeh, M. R., Gheisari, M., Liu, P., & Ashktorab, N. (2019). Hybrid model of ivfrn-bwm and robust goal programming in agile and flexible supply chain, a case study: Automobile industry. *IEEE Access, 7*, 71481–71492.

Henriques, C. O., Neves, M. E., Castelão, L., & Nguyen, D. K. (2022). Assessing the performance of exchange traded funds in the energy sector: A hybrid DEA multiobjective linear programming approach. *Annals of Operations Research, 313*(1), 341–366.

Hosseinzadeh, M. M., Ortobelli Lozza, S., Hosseinzadeh Lotfi, F., & Moriggia, V. (2023). Portfolio optimization with asset preselection using data envelopment analysis. *Central European Journal of Operations Research, 31*(1), 287–310.

Hsu, C. M. (2014). An integrated portfolio optimisation procedure based on data envelopment analysis, artificial bee colony algorithm and genetic programming. *International Journal of Systems Science, 45*(12), 2645–2664.

Huang, C. Y., Chiou, C. C., Wu, T. H., & Yang, S. C. (2015). An integrated DEA-MODM methodology for portfolio optimization. *Operational Research, 15*, 115–136.

Jatuphatwarodom, N., Jones, D. F., & Ouelhadj, D. (2018). A mixed-model multi-objective analysis of strategic supply chain decision support in the Thai silk industry. *Annals of Operations Research, 267*(1), 221–247.

Jones, D., & Jimenez, M. (2013). Incorporating additional meta-objectives into the extended lexicographic goal programming framework. *European Journal of Operational Research, 227*(2), 343–349.

Jones, D., & Tamiz, M. (2010). *Practical goal programming* (Vol. 141). Springer.

Jones, D., Firouzy, S., Labib, A., & Argyriou, A. V. (2022). Multiple criteria model for allocating new medical robotic devices to treatment centres. *European Journal of Operational Research, 297*(2), 652–664.

Kim, J. H., Lee, Y., Kim, W. C., & Fabozzi, F. J. (2022). Goal-based investing based on multi-stage robust portfolio optimization. *Annals of Operations Research*, pp. 1–18.

Kuchta, D. (2004). Robust goal programming. *Control and Cybernetics, 33*, 501–510.

Kuo, K. C., Lu, W. M., & Dinh, T. N. (2021). An integrated efficiency evaluation of China stock market. *Journal of the Operational Research Society, 72*(4), 950–969.

Lee, S. M., & Chesser, D. L. (1980). Goal programming for portfolio selection. *The Journal of Portfolio Management, 6*(3), 22–26.

Li, Z., & Floudas, C. A. (2012). Robust counterpart optimization: Uncertainty sets, formulations and probabilistic guarantees. In *proceedings of the 6th conference on foundations of computer-aided process operations, Savannah (Georgia)*.

Li, Z., Ding, R., & Floudas, C. A. (2011). A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization. *Industrial & engineering chemistry research, 50*(18), 10567–10603.

Lim, S., Oh, K. W., & Zhu, J. (2014). Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. *European Journal of Operational Research, 236*(1), 361–368.

Lin, R., & Li, Z. (2020). Directional distance based diversification super-efficiency DEA models for mutual funds. *Omega, 97*, 102096.

Liu, W., Zhou, Z., Liu, D., & Xiao, H. (2015). Estimation of portfolio efficiency via DEA. *Omega, 52*, 107–118.

Mansour, N., Cherif, M. S., & Abdelfattah, W. (2019). Multi-objective imprecise programming for financial portfolio selection with fuzzy returns. *Expert Systems with Applications, 138*, 112810.

Markowitz, H. (1952). Portfolio Selection. *The. Journal of Finance, 7*(1), 77–91.

Mensah, E. K., & Rocca, M. (2019). Light robust goal programming. *Mathematical and Computational Applications, 24*(4), 85.

Messaoudi, L., Aouni, B., & Rebai, A. (2017). Fuzzy chance-constrained goal programming model for multi-attribute financial portfolio selection. *Annals of Operations Research, 251*(1–2), 193–204.

Namdari, A., & Li, Z. S. (2018, June). Integrating fundamental and technical analysis of stock market through multi-layer perceptron. In *2018 IEEE technology and engineering management conference (TEMSCON)* (pp. 1–6). IEEE.

Oliveira, W. A., Fiorotto, D. J., Song, X., & Jones, D. F. (2021). An extended goal programming model for the multiobjective integrated lot-sizing and cutting stock problem. *European Journal of Operational Research, 295*(3), 996–1007.

Parkin, D., & Hollingsworth, B. (1997). Measuring production efficiency of acute hospitals in Scotland, 1991–94: Validity issues in data envelopment analysis. *Applied Economics, 29*(11), 1425–1433.

Pätäri, E., Leivo, T., & Honkapuro, S. (2012). Enhancement of equity portfolio performance using data envelopment analysis. *European Journal of Operational Research, 220*(3), 786–797.

Peykani, P., Mohammadi, E., Jabbarzadeh, A., Rostamy-Malkhalifeh, M., & Pishvaee, M. S. (2020). A novel two-phase robust portfolio selection and optimization approach under uncertainty: A case study of Tehran stock exchange. *PLoS ONE, 15*(10), e0239810.

Peykani, P., Seyed Esmaeili, F. S., Mirmozaffari, M., Jabbarzadeh, A., & Khamechian, M. (2022). Input/Output Variables Selection in Data Envelopment Analysis: A Shannon Entropy Approach. *Machine Learning and Knowledge Extraction, 4*(3), 688–699.

Rahiminezhad Galankashi, M., Mokhatab Rafiei, F., & Ghezelbash, M. (2020). Portfolio selection: A fuzzy-ANP approach. *Financial Innovation, 6*(1), 1–34.

Repetto, M., La Torre, D., & Liuzzi, D. (2019, April). A robust goal programming model for transfer pricing risk hedging: Preliminary results. In *2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)* (pp. 1–5). IEEE.

Romero, C. (2001). Extended lexicographic goal programming: A unifying approach. *Omega, 29*(1), 63–71.

Romero, C. (2004). A general structure of achievement function for a goal programming model. *European Journal of Operational Research, 153*(3), 675–686.

Sadjadi, S. J., Gharakhani, M., & Safari, E. (2012). Robust optimization framework for cardinality constrained portfolio problem. *Applied Soft Computing, 12*(1), 91–99.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119–138.

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research, 21*(5), 1154–1157.

Spronk, J., Steuer, R. E., & Zopounidis, C. (2016). Multicriteria decision aid/analysis in finance. In *Multiple Criteria Decision Analysis* (pp. 1011–1065). Springer, New York, NY.

Steuer, R. E., Qi, Y., & Hirschberger, M. (2007). Suitable-portfolio investors, nondominated frontier sensitivity, and the effect of multiple objectives on standard portfolio selection. *Annals of Operations Research, 152*(1), 297–317.

Steuer, R. E., Qi, Y., & Hirschberger, M. (2008). Portfolio selection in the presence of multiple criteria. In *Handbook of financial engineering* (pp. 3–24). Springer, Boston, MA.

Stoyan, S. J., & Kwon, R. H. (2011). A stochastic-goal mixed-integer programming approach for integrated stock and bond portfolio optimization. *Computers and Industrial Engineering, 61*(4), 1285–1295.

Tamiz, M., & Azmi, R. A. (2019). Goal programming with extended factors for portfolio selection. *International Transactions in Operational Research, 26*(6), 2324–2336.

Tamiz, M., Azmi, R. A., & Jones, D. F. (2013). On selecting portfolio of international mutual funds using goal programming with extended factors. *European Journal of Operational Research, 226*(3), 560–576.

Tarnaud, A. C., & Leleu, H. (2018). Portfolio analysis with DEA: Prior to choosing a model. *Omega, 75*, 57–76.

Tsolas, I. E. (2022). Performance evaluation of utility exchange-traded funds: A super-efficiency approach. *Journal of Risk and Financial Management, 15*(7), 318.

Uría, M. V. R., Caballero, R., Ruiz, F., & Romero, C. (2002). Meta-goal programming. *European Journal of Operational Research, 136*(2), 422–429.

Wang, L., & Li, J. (2019, August). A robust weighted goal programming approach for supplier selection problem with inventory management and vehicle allocation in uncertain environment. In *International Conference on Management Science and Engineering Management* (pp. 295–309). Springer, Cham.

Wu, Q., Liu, X., Qin, J., Zhou, L., Mardani, A., & Deveci, M. (2022). An integrated generalized TODIM model for portfolio selection based on financial performance of firms. *Knowledge-Based Systems, 249*, 108794.

Xidonas, P., Mavrotas, G., & Psarras, J. (2009). A multicriteria methodology for equity selection using financial analysis. *Computers and Operations Research, 36*(12), 3187–3203.

Xidonas, P., Steuer, R., & Hassapis, C. (2020). Robust portfolio optimization: A categorized bibliographic review. *Annals of Operations Research, 292*(1), 533–552.

Xidonas, P., Mavrotas, G., Krintas, T., Psarras, J., & Zopounidis, C. (2012). Multicriteria portfolio management. In *Multicriteria Portfolio Management* (pp. 5–21). Springer, New York, NY.

Yu, J. R., & Lee, W. Y. (2011). Portfolio rebalancing model using multiple criteria. *European Journal of Operational Research, 209*(2), 166–175.

Zhou, Z., Gao, M., Xiao, H., Wang, R., & Liu, W. (2021). Big data and portfolio optimization: A novel approach integrating DEA with multiple data sources. *Omega, 104*, 102479.

Zhu, J. (2022). DEA under big data: Data enabled analytics and network data envelopment analysis. *Annals of Operations Research, 309*(2), 761–783.

Zopounidis, C., Galariotis, E., Doumpos, M., Sarri, S., & AndriosopouloS, K. (2015). Multiple criteria decision aiding for finance: An updated bibliographic survey. *European Journal of Operational Research, 247*(2), 339–348.

Authors and Affiliations

Naeem Mohseny-Tonekabony[1] · Seyed Jafar Sadjadi[1] · Emran Mohammadi[1] · Mehrdad Tamiz[2] · Dylan F. Jones[3]

✉ Dylan F. Jones
  dylan.jones@port.ac.uk

  Naeem Mohseny-Tonekabony
  naeem_mohseny@ind.iust.ac.ir

  Seyed Jafar Sadjadi
  sjsadjadi@iust.ac.ir

  Emran Mohammadi
  e_mohammadi@iust.ac.ir

  Mehrdad Tamiz
  mehrdad.tamiz@yasar.edu.tr

[1] School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

[2] Department of Industrial Engineering, Yaşar University, Izmir, Turkey

[3] Centre for Operational Research and Logistics, School of Mathematics and Physics, University of Portsmouth, Portsmouth, UK