by

SUPPLY CHAIN ANALYSIS IN MINERAL AND ENERGY MARKETS

by

Merve Olmez Turan

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Mineral and Energy Economics).

Golden, Colorado

Date ______________________

Signed: ______________________

Merve Olmez Turan

Signed: ______________________

Dr. Tulay Flamand
Thesis Advisor

Golden, Colorado

Date ______________________

Signed: ______________________

Dr. Scott Houser
Department Head
Department of Economics and Business

ABSTRACT

This thesis presents a series of studies on supply chain analysis in mineral and energy markets. These studies include analyses of decisions on lithium and natural gas supply chains, demand-side measures to understand the natural gas market fragility, and effects of carbon policy on lithium production. Chapter 2 compares two commonly applied methodologies for estimating the causal effect of supply shifters on natural gas demand which are instrumental variables (IV), with variations of post-LASSO applications. Also, this chapter provides residential price elasticities for 23 European Union (EU) countries. Chapter 3 proposes a mixed-integer programming model in order to prescribe the optimal infrastructural decisions for the EU natural gas market that maximizes consumers' welfare. Chapter 4 examines how to supply lithium demand for the next 20 years by proposing a mixed-integer programming model in order to provide optimal decisions with respect to economical and environmental factors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

During the last four years, I have accomplished countless skills in the Department of Economics and Business and other departments at the Colorado School of Mines. I would like to present my truthful gratitude and appreciation to my supervisor Dr. Tulay Flamand for her marvelous mentorship. My achievements in the program would not have been the same without her guidance. I am also grateful to Dr. Roderick Eggert and Dr. Ben Gilbert for their valuable advice and support along the way. Lastly, I would like to offer my special thanks to Dr. Alexandra Newman and Dr. Max Brown for their insightful comments and suggestions.

This venture would not have been possible without the support of the Department of Energy's Critical Materials Institute. Through their generosity, I have been able to pursue purposeful projects as a part of the amazing research team. Also, I would like to thank VaasaETT, specifically Dr. Philip Lewis and Rafaila Grigoriou, for supporting our study by providing monthly residential price data for 23 EU countries.

I thank my friends, Sayeh Shojaeinia, Basar Basbug, Gamze Kalay, Gurbet Gurkan, Cansu, and Cagdas Demirkan, for their endless support. Whenever I need them, they always are with me. I would like to thank my research team, Sangita Gayatri Kannan, Maxwell Fleming, Shivani Mathur, and Jusse ALine Hirwa, for a cherished time spent together in the department and social settings.

I would like to express my special thanks to my parents, Nuray and Huseyin Olmez, and my sisters, Esma Sultan Yıldırım and Nihle Bozdag, for their encouragement and belief in me. Finally, I owe my biggest appreciation to my husband, Zeki, for his countless sacrifices and efforts. And my little son, Kerem, always enlightens my life with his smile.

Dedicated to my mother.

CHAPTER 1

INTRODUCTION

Many countries' governments aim to reduce the effect of their growth in manufacturing on the environment by decreasing carbon emissions. To achieve their target, they have an acting plan for energy transactions from high-carbon energy to clean energy sources. These plans can be investigated in two groups ,namely, share of final energy consumption and mineral markets for clean-energy technologies. Also, the policymakers make sure that their decisions consider supply security and demand estimation, where supply chain analysis covers both of these criteria. In this dissertation, we examine the natural gas supply chain for the European Union (EU), which is the most effective energy source for clean energy transactions in the short to medium-term. We also examine the global closed-loop lithium supply chain.

In Chapter 2, we estimate the price elasticity of residential natural gas demand for 23 EU countries using monthly data from 2014 to 2019. In the literature, neighboring countries' weather shocks can act as a supply shifter to identify demand. However, it is unclear which subset of neighbors in practice should be used. To address this issue, we employ and compare four traditional instrumental variables (IV) models to several post-LASSO approaches: post-LASSO OLS, post-LASSO IV, and two-stage post-LASSO IV at the country and panel-level. Our key contributions to the literature are the comparison of post-LASSO IV and two-stage post-LASSO IV with variations of the traditional IV method both at the country-level and panel-level. Moreover, we provide and discuss the causal effects of the supply shifters on individual residential natural gas price elasticities for a large set of EU countries and a single elasticity for the 23-country panel.

In Chapter 3, we examine the European natural gas supply chain with several tiers, including producers, mid-streamers, and consumers. We also consider that natural gas and Liquefied Natural Gas (LNG) could be traded via long-term contracts or spot markets. This network problem is formulated as a non-linear mixed-integer programming model which prescribes the optimal production and export decisions of producers, import and storage decisions of mid-streamers, and infrastructure decisions in order to maximize the total social welfare in the European Union (EU) over a five-year horizon. The proposed model seeks to answer the following questions: (i) What are the optimal production and export decisions of producers, import and storage decisions of mid-streamers, and infrastructure decisions of the regulator that jointly maximize the social welfare in the European natural gas market? (ii) What is the effect of these investment decisions on the social welfare? (iii) How much do these investment decisions change the market share of producers and market structure? (iv) How the social welfare changes if Russia was not considered as a supplier in the EU natural gas market? To the best of our knowledge, this is the first mixed-integer non-linear programming formulation that jointly considers monthly trade and storage decisions of all supply chain components via long-term contracts or spot markets, as well as infrastructure decisions of opening new pipelines and LNG regasification terminals. Moreover, this is the first study examining the effect of these infrastructure decisions on the supply risk and the effect of Russian gas interruption on the EU natural gas market.

In Chapter 4, we investigate the closed-loop lithium supply chain with several actors, including primary and secondary raw material producers (RMPs) (i.e., brine, hard rock mines, recycling facilities), mineral conversion plants, cathode manufacturers, battery and final product manufacturers, end-of-life (EOL) battery storage, and end-users in order to analyze future supplier dynamics in the market. We propose a mixed-integer linear programming model in order to determine the optimal amounts of lithium-carbonate traded among these supply chain actors, annual capacities for primary RMPs as well as investment decisions for new brine and hard rock deposits in order to minimize the total cost, which includes direct cost and cost of carbon from 2019 to 2040. In addition, using the proposed model, we examine the optimal decisions to answer the questions: (i) What are the optimal decisions for the selection of primary and secondary RMPs, their annual capacities and production levels, and the amount of lithium transported between all supply chain components that minimize the total cost (i.e., capital costs, operational and transportation costs and the cost of carbon) while meeting the future lithium requirement? (ii) How sensitive are these decisions and the global warming potential (GWP) of the supply chain to the changes in the social cost of carbon (SCC)? (iii) How do additional cathode manufacturers in the United States (US) and Europe affect the extraction and recycling decisions of the lithium supply chain? Our contributions are that this is the first optimization model that represents the global closed-loop lithium supply chain in order to prescribe the share of RMPs and the effect of the SCC and cathode manufacturer's location on the share of RMPs, country-level market concentration in extraction stage, and the GWP of the supply chain.

The remainder of this dissertation is organized as follows: Chapter 2 provides estimates for 23 EU countries' residential natural gas demand by using two different approaches, traditional IV and post-LASSO IV. Moreover, it also compares these methodologies with respect to three criteria performance. Chapter 3 prescribes optimal infrastructural investment decisions for the European natural gas supply chain problem in order to maximize social welfare. This chapter also discusses the effect of these investments on the markets under two scenarios, business-as-usual and Russian gas interruption. Chapter 4 studies the global closed-loop lithium supply chain in order to prescribe raw material producers' shares and discuss the effect of the SCC on these shares and GWP for the supply chain for the next 20 years. Finally, Chapter 5 concludes the research with a summary of our findings and directions for future research.

CHAPTER 2

HOW GOOD ARE WEATHER SHOCKS FOR IDENTIFYING ENERGY ELASTICITIES? A LASSO-IV APPROACH TO EUROPEAN NATURAL GAS DEMAND

Submitted to *The Journal of the Association of Environmental and Resource Economics.*

Merve Olmez Turan[1,2], Benjamin Gilbert[3], Tulay Flamand[3]

2.1 Abstract

We estimate the price elasticity of residential natural gas demand for 23 European Union (EU) countries using monthly data from 2014 to 2019. While neighboring countries' weather shocks can in theory act as a supply shifter to identify demand, it is unclear which subset of neighbors in practice should be used, if any. To address this issue, we compare four traditional instrumental variables (IV) models to several post-LASSO approaches: post-LASSO OLS, post-LASSO IV, and two-stage post-LASSO IV. We compare these models on a country-by-country basis and for the full panel. We find that two-stage post-LASSO IV performs best in most cases for individual countries. It also has the most reliable results at the panel-level. Our preferred estimates suggest that country-level price elasticities range from −0.98 to −0.09, with a median of −0.58, in line with estimates from previous literature.

2.2 Introduction

Natural gas is the main energy source for residential consumers in the European Union (EU). European natural gas consumption has major geopolitical and environmental implications. Russia is responsible for 40% of imported gas to Europe, making it the dominant supplier in the region [1]. Russia's invasion of Ukraine has raised serious questions about the impact on EU households if Europe reduces natural gas imports from Russia, or Russia reduces exports to Europe [2]. Longer term efforts by Europe to decarbonize its economy have implied an increasing reliance on natural gas, whereas more recent efforts to reduce methane emissions from the supply chain may increase costs of production and transportation [3, 4]. The demand elasticity is a key parameter for understanding how the EU economy will fare under natural gas price increases.

Our goals in this paper are twofold: to provide reliable European residential natural gas price elasticity estimates, and to demonstrate whether recent advances in causal methods using machine learning can improve upon traditional instrumental variable (IV) methods for elasticity estimation in energy markets.

[1]Graduate student, Colorado School of Mines
[2]Primary researcher and author
[3] Assistant Professor, Colorado School of Mines

Using a variety of traditional IV and post-LASSO IV models, we estimate country-level residential natural gas price elasticities for 23 European Union countries, as well as a single elasticity for the full panel of countries in our data set. Reliable estimates of these elasticities will help policymakers in order to understand how householders will react to price changes in EU natural gas markets. Yet there are relatively few studies in the literature with updated, country-specific estimates of residential natural gas price elasticities in Europe. Exceptions include Germany [5], the Netherlands [6], the United Kingdom [7], and Switzerland [8]. To the best of our knowledge, this is the first study to estimate residential natural gas price elasticities for such a large number of individual EU countries using a consistent dataset and methods. Moreover, our results provide guidance on methodological tradeoffs in elasticity estimation for energy commodities.

In order to arrive at our residential natural gas price elasticity estimates, we compare four traditional IV specifications to post-LASSO Ordinary Least Squares, post-LASSO IV, and two-stage post-LASSO IV methods at both the country- and panel-level. In order to address price endogeneity, we consider other countries' weather shocks, specifically cooling degree days (CDD) and heating degree days (HDD) as exogenous supply shifters for a given country. As weather-driven demand in neighboring countries varies, the supply available to a given country also changes. Using neighboring regions' weather as an IV for demand has been used previously in the United States (US) context by, for example, Hausman and Kellogg [9], and Davis and Muehlegger [10]. It is not ex ante obvious, however, which neighboring region's weather shocks are the most relevant supply shifters for a given country, if any. Identifying the optimal set of IVs also poses challenges, particularly if the number of candidate instrumental variables is large relative to the total number of observations - as is the case in our study with only five years of monthly observations on each country. We therefore employ the post-LASSO IV methodology, proposed by Belloni et al. [11] to tackle these issues. In an initial stage, post-LASSO IV uses LASSO to select the optimal set of instrumental variables among all given variables before implementing the standard IV estimator on the selected set. We also propose and execute a two-stage post-LASSO IV method by inserting an additional initial LASSO step that selects the optimal set of control variables before selecting the optimal instruments, and finally implements the standard IV estimator using the selected controls and IVs. In order to demonstrate the importance of price endogeneity for elasticity estimates, we also compare results from these models to post-LASSO OLS.

We then evaluate these models according to three criteria: (i) reasonableness: the estimated natural gas price elasticity should be in the (-1,0) interval, (ii) relevance: the selected IVs should produce a first-stage F-statistic greater than 10, and (iii) validity: if more than one IV is selected, then the IVs should pass an overidentification test. The "reasonableness" criterion is motivated by the vast majority of previous studies which find that short-run residential natural gas demand is inelastic. We review this literature more extensively in the next section. We therefore assume that short run elasticity estimates that fall outside the (-1,0) interval are likely driven by some combination of endogenous regressors, weak instruments, and sampling variation.

We find that at the country level, two-stage post-LASSO IV outperforms the other models by satisfying the greatest number of our three criteria for the greatest number of countries. However, other models perform systematically better along some dimensions and worse along others. For example, using just the four nearest countries' HDDs and CDDs as instrumental variables satisfies the relevance criterion for the greatest number of countries. Moreover, some of our traditional IV models satisfy the validity criterion for all 23 EU countries. Two-stage post-LASSO IV satisfies the "reasonableness" criterion for the greatest number of countries, but it fulfills the "validity" criterion for the lowest number of countries, 16, providing invalid instruments for the remaining seven countries. This result supports the study of Kang et al. [12] that suggests that the IV selection of LASSO may include invalid instruments in the "optimal" selected set. On the other hand, the two-stage post-LASSO IV model yields better estimates at the panel level in terms of satisfying the first and second criteria. The third criterion is not applicable at the panel level because post-LASSO IV selected only one instrumental variable and the overidentification test cannot be performed.

We also find that two of our traditional IV models for Poland, and the two-stage post-LASSO IV model for Portugal and Slovakia satisfy all three criteria, implying that elasticities obtained from these models for these countries are highly reliable. On the other hand, none of the models for Austria, Latvia, and Slovenia satisfy either the first or third criteria. We therefore conclude that results for these countries are unreliable. For the other 17 EU countries, none of the models satisfy the second criterion (relevance) insofar as their first-stage F-statistics do not exceed the rule-of-thumb value of 10. To this end, for each country, we consider the model that satisfies the first and third criteria, while providing the highest first-stage F-statistic as the "preferred" model, even if that value does not exceed 10. Our preferred estimates of residential natural gas price elasticities for 20 EU countries are found to be in the acceptable range with a median of -0.58. Moreover, the two-stage post-LASSO IV model at the panel-level estimates residential natural gas price elasticity as -0.32 for these 23 EU countries as a whole.

In the literature, several papers have estimated the price elasticity for the overall natural gas market, while there are fewer studies focusing on estimating the natural gas price elasticity specifically for the residential sector in individual European countries [5–7]. Our work is most similar in this regard to Asche et al. [13] who estimate individual residential natural gas demand elasticities for 12 European countries using the panel shrinkage estimator of Maddala et al. [14]. While their shrinkage estimator provides a

[3] Much of the literature also finds that long-run residential natural gas demand is also inelastic or near inelastic.

compromise between country-by-country OLS versus panel aggregate estimation in terms of the number of parameters to be estimated, they do not isolate a source of exogenous variation in price in order to identify the demand curve. By contrast, we use LASSO – which has also been referred to as a shrinkage estimator – to "shrink" the set of available instruments to the best exogenous supply shifters for each country. We also provide elasticity estimates for additional countries, using more recent data. As mentioned, our IV strategy is most closely related to Hausman and Kellogg [9] and Davis and Muehlegger [10] in terms of using neighboring regions' weather shocks as IVs. However, these studies use aggregated weather shocks as IVs. Aggregating strong and weak variables may diminish the effect of strong instruments, and therefore selecting the optimal set of variables by eliminating weak ones may improve the model. Our key contributions are in (i) comparing post-LASSO IV and two-stage post-LASSO IV with variations of the traditional IV method both at the country level and panel level, and (ii) providing and discussing individual residential natural gas price elasticities for a large set of EU countries, and a single elasticity for the 23-country panel.

The remainder of this paper is organized as follows: Section 2.3 provides a detailed literature review. Section 2.4 describes the empirical models and Section 2.5 provides details regarding the data acquisition. Section 2.6 discusses our empirical results regarding the comparison of models' performance. This section also discusses insights based on reliable residential price elasticities at country- and panel-level. Finally, Section 2.7 concludes the paper with a discussion of future work.

2.3 Literature Review

In the literature, there are many studies that estimate the natural gas price elasticity for EU countries at the panel- and/or country-level. Most of these find that demand is inelastic. One of the most comprehensive studies on the residential natural gas demand for 12 European countries at the individual country and panel levels is by Asche et al. [13]. They estimated the association between residential natural gas consumption per capita and residential natural gas price, residential light fuel oil price, residential electrical price, and personal income per capita from 1960 to 2002 using a dynamic log-linear model. Asche et al. [13] suggest that at the country-level, the short-run and long-run price elasticities of natural gas demand are in the range of 0.0 to -0.3, and 0.0 to -0.6, respectively, while the panel-level analysis yields short-run and long-run price elasticities of -0.24 and -1.54, respectively. At the panel-level, [15] examines 23 of the Organization for Economic Co-operation and Development (OECD) countries during the time period between 1978 and 1999 and shows that the long-run price elasticity of residential natural gas demand for these countries is -0.25. Another study, however, reports the long-run residential price elasticity for 35 OECD countries and nine non-OECD countries to be -1.25 during the time period between 1978 and 2011

[16]. Moreover, during the same time period, the long-run natural gas price elasticity for specifically OECD-Europe countries is reported as -0.16 [17]. At the country-level, Bernstein and Madlener [18] investigate residential natural gas demand for ten OECD-Europe countries and two OECD countries using the autoregressive distributed lag bounds testing procedure for annual time-series data from 1980 to 2008. The countries covered in their study are Austria, Finland, France, Germany, Ireland, Japan, Luxembourg, the Netherlands, Spain, Switzerland, the UK, and the US. They suggest that the short-run and long-run average price elasticity are -0.24 and -0.51, respectively. Moreover, [6] estimate the residential price elasticity for Netherlands as -0.19, [5] estimates that for Germany in the range of -0.63 and -0.44, [7] calculate that for the United Kingdom in the range of -0.56 to -0.34, while the residential price elasticity for Switzerland is estimated to be -0.73 [19] and for Greece is estimated in the range of -0.90 to -0.82 [20]. Note that, all aforementioned studies that estimate the residential natural gas demand for European countries employ several methodologies such as the structural time series model [17], random effect model [7], or quadratic expenditure system [5]. In addition, Filippini and Kumar [19] estimate the causal effect of natural gas price with supply shifters for Switzerland, and Burke and Yang [16] estimate that for 44 OECD and non-OECD countries at panel-level. Specifically, Filippini and Kumar [19] use the mean of the average prices faced by all other local dwellings, and Burke and Yang [16] use per capita domestic natural gas reserves and negative power distance-weighted reserves in other countries as supply-side IV in their models.

In order to identify elasticities of demand, the resource economics literature contains a number of studies at either the panel or country level that use various supply-side IV in order to address the potential for endogeneity in natural gas prices [9, 10, 21]. Specifically, [10] estimate natural gas demand for the residential, commercial, and industrial sectors, in the United States. In order to estimate the natural gas demand of a state in the commercial and residential sectors, the authors consider HDDs and CDDs of all other states as an instrument for the natural gas price. Neighboring regions' weather is certainly not the only supply shifter for resource commodities; for example, Roberts and Schlenker [22] use past storage shocks as an instrument for identifying supply elasticities for agricultural commodities, which are storable natural resources. Finally, to estimate the natural gas demand of a state, Hausman and Kellogg [9] combine the instruments used in [10] and Roberts and Schlenker [22] in order to propose a new instrument that is the sum of all other states' population-weighted average of HDDs over the second through the 12th monthly lag. In our study, we propose four IV models where the instrumental variables are either: (i) the sum of all other countries' weather shocks, or (ii) the four nearest countries' weather shocks included individually. In two of these models, the 12-month lag of natural gas consumption is included as a control variables.

Recall that in order to estimate the natural gas demand of a state or country, the aforementioned studies use the aggregate effects of other regions' weather and/or inventory shocks as instruments. However, some of the components of these aggregate instrumental variables may be stronger price shifters than others, and aggregating these strong and weak variables may reduce the effect of more exogenous ones on the model, and reduce the accuracy of the first-stage prediction. Including all components of the aggregate instrument separately, however, may result in a first stage with many weak instruments. It may also be infeasible to consider many individual instrumental variables simultaneously in the model when the number of instrumental variables is large relative to the number of observations. To overcome this challenge, Belloni et al. [11] develop a post-LASSO IV that uses LASSO regression in an initial step of the first-stage model in order to select the best subset of all available instrumental variables, before using the selected instruments in the standard IV estimator. This procedure reduces the number of instrumental variables in the model to those that are the strongest predictors of the endogenous variable. Belloni et al. [23] propose sets of sparse methods such as LASSO, Post-LASSO, $\sqrt{LASSO}$, and $\sqrt{Post-LASSO}$. Comparing these estimators with the IV estimator which considers all possible instrumental variables, suggests that these sparse models with data-driven penalties perform better than the alternative with many instruments in terms of mean absolute deviation and root-mean-squared error. LASSO IV methods have been applied to study the estimation of demand for automobiles [24], the estimation of price elasticities of groceries [25], the analyse of gene expression and genetic variants [26], the effect of body mass index on diastolic blood pressure [27], and the effect of abortion on crime rates [28]. To the best of our knowledge, our study is the first to apply the post-LASSO IV method in order to estimate natural gas demand elasticity.

2.4 Empirical Models

In this section, we propose four methods, namely, post-LASSO OLS, IV, two-stage post-LASSO IV, and post-LASSO IV in order to estimate the residential natural gas price elasticities.

The LASSO method [29] aims to fit a function that minimizes the sum of square errors, plus an additional penalty for the sum of the absolute values of the coefficients in order to determine the strongest variables for the model. While LASSO itself provides a predictive model with these variables, "post-LASSO" OLS, proposed by Belloni and Chernozhukov [30] performs OLS on the subset of selected independent variables. However, to identify the demand elasticity of a market where supply also depends on the market price, the OLS estimator may be biased due to the endogeneity of the price. IV can be used to deal with these issues, however, the application of the IV method may be challenging to find strong instrumental variables. Therefore, in order to select the optimal set of instrumental variables, Belloni et al. [11] proposed the post-LASSO IV methodology. We further suggest a two-stage post-LASSO IV method to select both the instrumental variables and control variables.

Our base specification is a standard demand function that regresses the natural log of a country's residential natural gas consumption $log(Q_{ct})$ on the natural log of its residential natural gas price $log(P_{ct})$, along with monthly dummy variables S_m that represent seasonal effects and a set of additional demand shifters $\mathcal{X}$, as shown in Equation 2.1:

$$\log(Q_{ct}) = \beta_0 + \beta_1 \log(P_{ct}) + \beta_i \mathcal{X} + \beta_m S_m + \varepsilon_{ct} \tag{2.1}$$

In our study, $\mathcal{X}$ may include a country's Heating Degree Days (HDD), Cooling Degree Days (CDD), Gross Domestic Production (GDP), and unemployment rate. We include $log(P_{ct})$ and S_m in every specification.

We first use post-LASSO OLS to (i) select the optimal set of demand shifters from $\mathcal{X}$, and (ii) fit Equation 2.1 using the selected regressors.

Next, we estimate four traditional IV specifications represented in Equation 2.2:

$$\begin{aligned} \log(Q_{ct}) &= \beta_0 + \beta_1 \widehat{\log(P_{ct})} + \beta_i \mathcal{Y} + \beta_m S_m + \varepsilon_{ct} \\ \widehat{\log(P_{ct})} &= \delta_i \mathcal{Z} \end{aligned} \tag{2.2}$$

Here, $\mathcal{Y}$ represents researcher-selected demand shifters, and $\mathcal{Z}$ represents the set of researcher-selected instrumental variables. In this paper, we examine four IV models, each of which includes different combinations of control and instrumental variables as shown in Table 2.1.

All specifications include HDD, CDD and monthly dummies in their set of control variables, and include European Brent Spot Crude Oil Price (EBSP) in their set of instrumental variables. Additionally, the sum of all other countries' HDD and CDD are used as instrumental variables in IV Model 1 and IV Model 2. In contrast to IV Model 1, IV Model 2 also includes the 12 month lag of the natural log of natural gas consumption as a control variable. Note that while IV Models 1 and 2 use the sum of all other countries' HDD, CDD, and EBSP, IV Models 3 and 4 use the four nearest countries' HDDs, CDDs, and EBSP instead, as instruments. IV Model 4 also includes the 12-month lag of logged consumption as a control variable. The four nearest countries are determined based on the distance between their capital cities.

Table 2.1 Control and instrumental variables for our IV models

Model	**$\mathcal{Y}$-Control Variables**	**$\mathcal{Z}$-Instrumental Variables**
IV Model 1	HDD_{ct}, CDD_{ct}	Sum of all other countries' HDD, CDD, and EBSP
IV Model 2	HDD_{ct}, CDD_{ct}, $\log(Q_{c(t-12)})$	Sum of all other countries' HDD, CDD, and EBSP
IV Model 3	HDD_{ct}, CDD_{ct}	Four nearest countries' HDDs, CDDs, and EBSP
IV Model 4	HDD_{ct}, CDD_{ct}, $\log(Q_{c(t-12)})$	Four nearest countries' HDDs, CDDs, and EBSP

[3]This is similar to the Davis and Muehlegger [10] study of US natural gas demand, however those authors also include, and instrument for, the interaction of price with each state's HDD and CDD which we omit for simplicity.

Finally, we discuss the post-LASSO IV and two-stage post-LASSO IV methods. While post-LASSO IV can determine the optimal set of instrumental variables, we suggest two-stage post-LASSO IV to additionally determine the optimal set of control variables. For each country, we consider the full set of all other 22 countries' HDDs and CDDs as potential instrumental variables, and use post-LASSO IV to select the optimal set of instruments among them. Moreover, we consider the country's own HDD, CDD, unemployment rate, and GDP as potential control variables, where two-stage post-LASSO IV employs an initial step on the second stage equation to select among these controls and uses the optimal set, along with monthly dummies. The two-stage post-LASSO IV methodology is illustrated in Figure 2.1. Specifically, this method includes five models. In the figure, the dashed and solid parallelograms represent the input of the model for country-level and panel-level analysis, respectively. Note that, depending on whether we perform this methodology at the panel-level or country-level, only one of these inputs are used in this methodology. In addition, large rectangles represent the models, while small rectangles highlighted in grey show their outputs. The first model shown in the figure uses a LASSO regression for Equation 2.1 on all potential control variables as well as the logarithmic natural gas price and monthly dummy variables in order to select the optimal set of controls. This set of optimal control variables is then used in the second, third, and fifth models. Specifically, the second model performs a linear regression of the logarithmic natural gas price on this optimal set of control variables and monthly dummies in order to residualize the treatment variable. The third model performs a set of linear models, one for each possible instrumental variable, on the optimal set of control variables and monthly dummies in order to residualize the set of possible instruments. The fourth model performs a LASSO regression of the residual natural gas price obtained by the second model on the set of residualized instrumental variables obtained by the third model in order to select the set of optimal instrumental variables to be used in the fifth model. Finally, the fifth model implements the standard IV estimator for Equation 2.2 using the set of optimal instrumental variables obtained by the fourth model, the set of optimal control variables obtained by the first model, the logarithmic natural gas price, and monthly dummies, in order to provide the residential natural gas price elasticity.

Lastly, as a final method to estimate the natural gas price elasticities, we perform the post-LASSO IV method proposed by Belloni et al. [11]. This method covers all steps of two-stage post-LASSO IV except for the first model, since it considers the set of control variables to be predeterimined without selecting them. Note that, in this study, HDD, CDD, and monthly dummy variables are given as the set of control variables, while the optimal set of instrumental variables are determined by selecting among other countries' HDDs and CDDs.

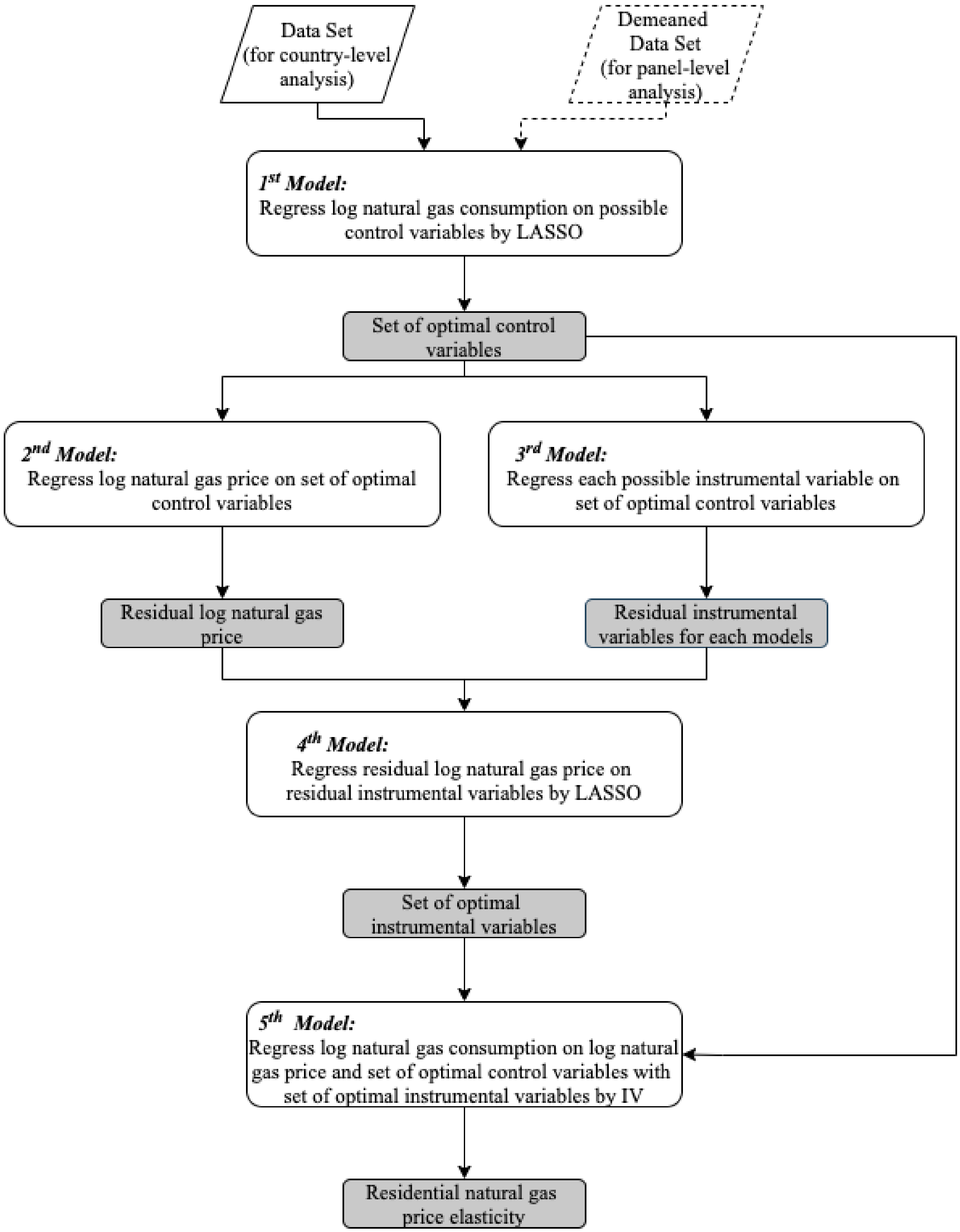


Figure 2.1 The two-stage post-LASSO IV methodology [11]. Dashed and solid parallelograms represent the input of the model for country-level and panel-level analysis, respectively. Large rectangles represent models, while small grey rectangles show their outputs

Although four methodologies are discussed above at the country-level, note that we perform all these methodologies for the price elasticity estimation also at the panel-level. To this end, we use the fixed-effect panel model, wherein demeaned variables are used instead of variables' real values.

2.5 Data Acquisition

Our data set includes monthly and quarterly data within a 60-month period from January 2014 to December 2018 for the 23 EU countries for which data was available, namely, Austria, Belgium, Bulgaria, Czechia, Denmark, Estonia, Greece, Spain, France, Croatia, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Romania, Sweden, Slovenia, and Slovakia. Natural gas is mainly used in the residential sector for space and water heating, so there exists monthly seasonality to consumption. In order to construct a monthly country-level panel, we collect data from several sources. First, we use the monthly residential natural gas price data collected by VaasaETT. This data provides the end-user price in each country's capital city, measured in Euro cents per kWh and including energy and value-added taxes. Note that, while in the country-level data, missing residential natural gas price observations for some countries in some months decreases sample sizes to about 48 to 60 observations, the panel-level data set includes 1357 observations for these 23 countries.

Because residential natural gas consumption is not publicly available at a monthly-level, we calculated monthly consumption by multiplying total monthly natural gas consumption from all sectors, obtained from the Eurostat Database, with the annual residential natural gas consumption share (out of the total sectoral consumption), obtained from the International Energy Agency (IEA).

Our data set also includes weather variables, which are the HDD and CDD of countries obtained from the Eurostat Database. We obtain data for Europe's Brent spot prices from the US Energy Information Administration (EIA). Quarterly data of GDP at market price and unemployment rates for countries are acquired from the Eurostat Database. A complete list of variable definitions and their corresponding sources is provided in Table 2.2.

2.6 Empirical Results

In this section, we compare and discuss the performance of the aforementioned models both at the country-level and panel-level. Based on the best model performed for each country, we also discuss their residential natural gas price elasticity.

2.6.1 Model Comparison

We first compare four IV, two-stage post-LASSO IV, and post-LASSO IV models' performances from country-by-country estimates, and then discuss these models at the panel-level. Recall that we assess these

Table 2.2 Variable definitions and data sources

Variable	**Definition**	**Unit**	**Source**
Residential natural gas consumption	Non available monthly residential natural gas consumption	Billion cubic meters (bcm)	Our calculation based on IEA and EuroStat
Residential natural gas price	Monthly residential natural gas price in the capital city of each country	Euro cents per kWh	VaasaETT
HDD	Monthly heating degree days degree	Celsius	EuroStat
CDD	Monthly cooling degree days degree	Celsius	EuroStat
EBSP	Europe brent spot crude oil price	Euro per barrel	EIA
Unemployment rate	Quarterly unemployment rate	Percentage	Eurostat
GDP	Quarterly GDP at market price	Million national currencies	EuroStat

models based on three criteria: (i) the estimated price elasticity should be within the interval of $[-1, 0]$, (ii) the F-statistics of the first-stage estimation of an IV model should be greater than 10 to ensure that instrumental variables are not weak, and (iii) the p-value of the Sargan test should be greater than 0.05 to ensure that instrumental variables of a model are valid. Note that the third criterion is examined only for the models where there are higher number of instrumental variables than that of endogenous variables. Note that post-LASSO OLS is not compared to the other models based on the last two criteria which apply to IV models. It is considered as our base model from which to compare the alternative models.

The post-LASSO OLS models at the country level yield elasticities within the range of $[-2.54, 0.38]$. The median of these estimated price elasticities is -0.34. Five out of 23 countries, namely Denmark, Hungary, Belgium, Sweden, and Slovakia, have their post-LASSO OLS price elasticity estimated out of the $[-1, 0]$ interval, so the median is calculated without taking into account the price elasticity of those countries. In addition, post-LASSO OLS at the panel-level estimates the price elasticity for all 23 countries as -0.46.

We compare the six IV models (traditional and post-LASSO) in Figure 2.2, which shows histograms of estimated price elasticities, first-stage F-statistics and Sargan test p-values, respectively, across the 23 countries. In the figure, bins that satisfy the associated performance criterion are highlighted in dark grey, those that do not satisfy the criterion are highlighted in light grey. In Appendix, the results of six models are provided for each country. Recall that the first criterion ensures that the price elasticity estimates should be within the range of $[-1.0, 0.0]$. The first column of Figure 2.2 includes histograms associated with this criterion. These histograms indicate that among the four traditional IV models, the elasticities from between four and nine of the 23 countries are in the acceptable range. By contrast, 15 and 17 of the 23 elasticities are in the acceptable range for two-stage post LASSO-IV and post-LASSO IV, respectively. Among these, two-stage post-LASSO IV estimates price elasticities in the acceptable range for the highest number of countries, while IV Model 2 meets this criterion for the lowest number of countries.

The second column of Figure 2.2 presents histograms of the first-stage F-statistics across countries. None of the first-stage results for IV Model 1, IV Model 2, or post-LASSO IV have first-stage F-statistics that are greater than 10. This may be a limitation of our sample, with only 48 to 60 observations on each country. However, this implies that the set of instrumental variables selected in these models may be considered as weak instruments. Moreover, the first-stage of IV Model 3, IV Model 4, and two-stage post-LASSO IV provide valid F-statistics for two, three, and two countries, with median values of F-statistics 2.42, 3.37, and 3.05, respectively. Results reveal that the IV Model 4, which jointly considers the four nearest countries' HDDs and CDDs as eight separate instruments, and controls for the lag 12 residential natural gas consumption, includes the strongest set of instrumental variables among those of other models with respect to their first-stage F-statistics.

Finally, histograms in the third column in Figure 2.2 represent results for the third criterion that ensures instrumental variables are valid with p-values of the Sargan test higher than 0.05. Results show that IV Model 1 and IV Model 2 pass the Sargan test for all 23 EU countries, while IV Model 3 and IV Model 4 pass that for 21 out of 23 countries. On the other hand, two-stage post-LASSO IV and post-LASSO IV models provide valid instrumental variables for 20 and 16 countries, respectively. These imply that IV Model 1 and IV Model 2 outperform other models with respect to the Sargan test. Kang et al. [12] suggest that LASSO-selected instrumental variables might be invalid while they are exceptionally strong predictors of the endogenous treatment variable. Comparison of these six models with respect to the third criterion does seem weakly consistent with the conclusion that post-LASSO IV may select more invalid instruments given the slightly higher rejection rate on the Sargan tests. However, these models do not produce unambiguously stronger first stages in our sample as can be seen by examining results for the relevance criterion.

Comparisons between models at the panel-level are shown in Table 2.3. The estimated price elasticity, first-stage F-statistics, and the p-values for the Sargan test are reported in the last three columns, respectively. Results show that all models except for IV Model 1 and post-LASSO IV satisfy the first criterion by providing a price elasticity in the range of $[-1.0, 0.0]$. In addition, only two-stage post-LASSO IV satisfies the second criterion by providing an F-statistics value for the first-stage that is greater than 10. This implies that the instrumental variables in all other model except for two-stage post-LASSO IV may be considered as weak instruments. Regarding the third criterion, IV Model 1 and IV Model 2 pass the

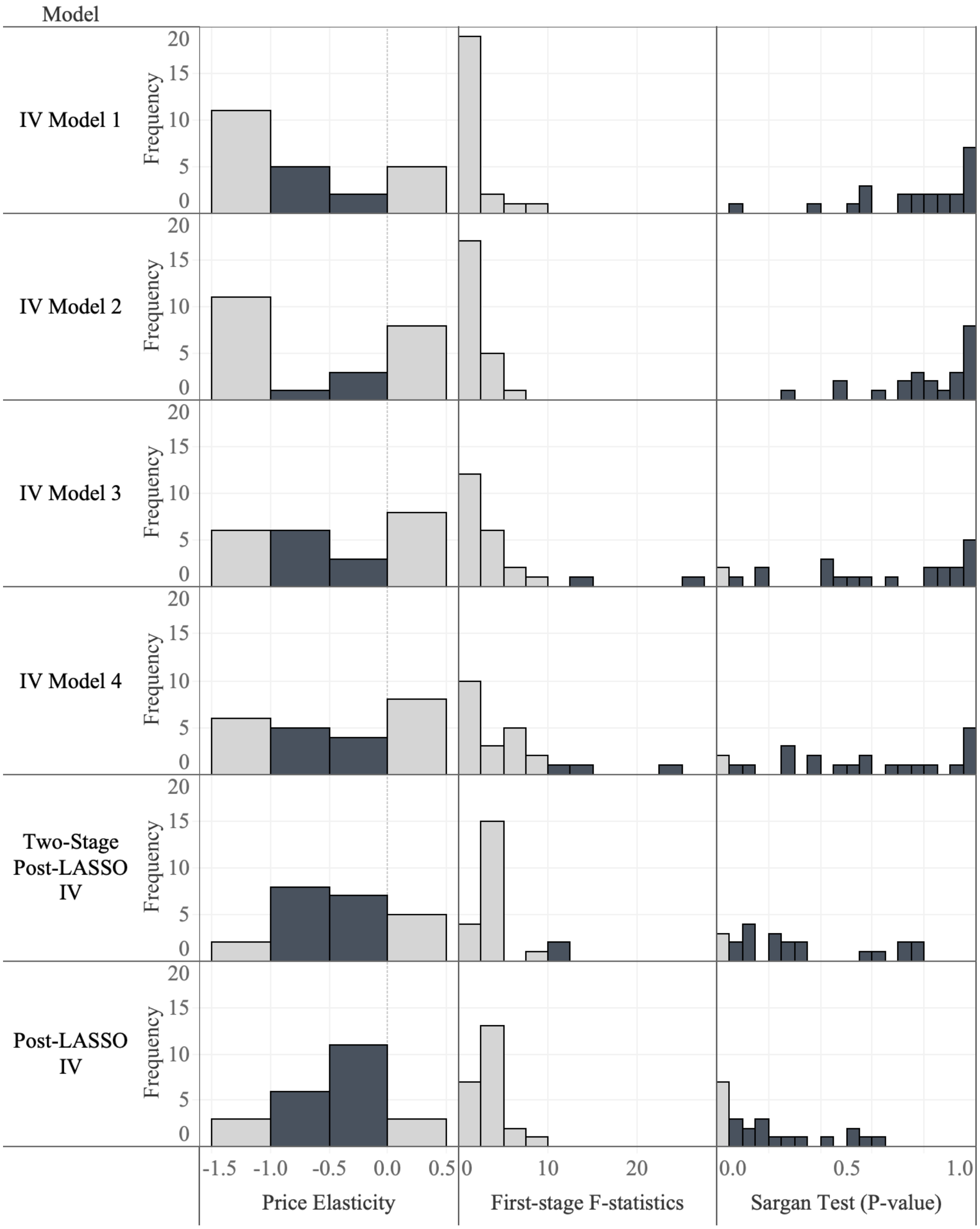


Figure 2.2 The histograms of price elasticity, first-stage F-statistics, and Sargan test p-value for four IV models, two-stage post-LASSO IV, and post-LASSO IV models at country-level

Sargan test by providing p-values more than 0.05, while IV Model 3, IV Model 4, and post-LASSO IV fail on this test. Note that the Sargan test is not applicable to the two-stage post-LASSO IV Model, since it selects only one instrument in this case. Overall, at the panel-level, this comparison reveals that the two-stage post-LASSO IV Model outperforms the other models by simultaneously satisfying the first and second criteria.

Table 2.3 Results of four IV models, two-stage post-LASSO IV, and post-LASSO IV models at panel-level

Empirical Models	**Instrumental Variables**	**Price Elasticity *(Std. Error)***	**First-stage F-statistics**	**Sargan Test *(p-value)***
IV Model 1	Sum of all other countries' HDD, CDD	-1.28 (*0.79*)	1.38	0.54
IV Model 2	Sum of all other countries' HDD and CDD	-0.57 (*0.37*)	2.76	0.43
IV Model 3	Four nearest countries' HDD and CDD	-0.84 (*0.84*)	1.32	≤ 0.05
IV Model 4	Four nearest countries' HDDs and CDDs	-0.43 (*0.47*)	1.59	≤ 0.05
Two-stage post-LASSO IV Model	Selected countries' HDDs and CDDs	-0.32 (*0.44*)	31.74	Not Applicable
Post-LASSO IV Model	Selected countries' HDDs and CDDs	0.61 (*0.24*)	8.44	≤ 0.05

2.6.2 Discussion on Residential Natural Gas Price Elasticities

In this section, we discuss our preferred price elasticity estimates obtained by the best performed model specific for each country, and the best panel model.

For the country-level analysis, we find that IV Models 3 and 4 for Poland, and the post-LASSO IV Model for Portugal and Slovakia, fulfill all three criteria. In addition, Austria, Latvia, and Slovenia fail on either the first or third criteria for all models. To this end, we do not discuss the price elasticities for these three countries, since none of the models can provide reliable estimates. As mentioned in Section 2.6.1, for the rest of the 17 EU countries, all models fail on the weak instruments criterion with F-statistics of less than 10 at their first-stage. Therefore, for each of these countries, the preferred model is selected among all six models such that it satisfies the first and the third criteria, and has the highest F-statistics (even if it does not satisfy the second criterion). Overall, we consider the estimated natural gas price elasticities obtained from IV Model 3 for four countries, IV Model 4 for three countries, post-LASSO IV for six countries, and two-stage post-LASSO IV for eight countries, which reveals that two-stage post-LASSO IV provides the preferred approach for the greatest number of countries.

Figure 2.3 presents a map of preferred residential natural gas price elasticity estimates for 20 EU countries (where Austria, Latvia, and Slovenia are not considered). In the figure, 20 EU countries are highlighted in gray scale, where darker shading represents more elastic countries. The selected price elasticities vary from -0.98 to -0.09. Although all are inelastic by construction, natural gas is most elastic for the residential market in Greece, France, and Italy, whereas it is most inelastic for Croatia and Lithuania compared to other countries. Moreover, the median of the preferred country-level price elasticities is -0.58, while the panel-level price elasticity is estimated as -0.32 by two-stage post-LASSO IV.

The estimated residential natural gas price elasticities for Poland, Portugal, and Slovakia, the three countries whose most preferred model meets all three criteria, are -0.36, -0.45, and -0.68, respectively. This implies that increasing the natural gas price by 100% causes a reduction in the residential natural gas consumption by 36%, 45%, and 68%, respectively for these three countries. These results suggest that Slovakia householders' natural gas consumption is more sensitive to a change in the natural gas price compared to those of Portugal and Poland householders. Note that, in EU countries, residential natural gas is mainly used for cooking, space and water heating, where for these purposes, electricity, renewable energy, delivered heat, and other fossil fuels (i.e., coal and oil) may also be used as alternative energy sources. Specifically, the transition from carbon intensive energy sources to renewable energy, may reduce dependency on natural gas. This substitutability may increase elasticity, since consumers may promptly react to price changes if they have several options for energy resources. Consistent with this observation, there was a 12% increase in the renewable energy consumption of Slovakian householders from 2013 to 2018, whereas in this time period, renewable energy use by households in Portugal and Poland declined by 3% and 9%, respectively (IRENA, 2021).

2.7 Conclusion

In this paper, we estimate and compare seven models to recover European residential natural gas demand elasticities. Namely, post-LASSO OLS, four traditional IV specifications (based on the consideration of different sets of instrumental and control variables), two-stage post-LASSO IV, and post-LASSO IV. We do so both at the country-level and panel-level in order to estimate the residential natural gas price elasticities for 23 EU countries. We compare the IV models based on three criteria that require (i) natural gas price elasticities to be in the range of -1.0 and 0.0, (ii) first-stage F-statistics to be greater than 10, and (iii) p-values of the Sargan test to be greater than 0.05. Comparison results indicate that the post-LASSO IV model consistently provides the most reliable estimates of residential natural gas price elasticity for the highest number of countries compared to other models at the individual

Figure 2.3 The map of preferred residential natural gas price elasticity for 20 EU countries

country-level. At the panel-level, two-stage post-LASSO IV model outperforms other models by satisfying the first and second criteria simultaneously. Note that, the third criterion is not applicable to two-stage post-LASSO IV since a single instrumental variable is selected by the LASSO routine.

By employing these models, we provide reliable residential natural gas price elasticity estimations for 20 out of 23 EU countries at country-level, while also estimate those at the panel-level based on the aggregated data over all these countries. The natural gas price elasticities have not been estimated in the literature by considering all 23 EU countries simultaneously at either the country- nor the panel-level. These price elasticity estimates play a significant role in order to provide insights to understand the residential market sensitivity to supply disruptions for the EU, which is a major natural gas importing region. In addition, while natural gas has the potential to be a low carbon bridge fuel, methane emissions from the natural gas supply chain have raised concerns about reliance on natural gas from a climate perspective.

Well-known papers in the literature use cumulative and neighboring weather shocks as supply shifters to isolate demand with in the traditional IV methodology. In repeating this exercise for the EU and comparing to two-stage post-LASSO IV and post-LASSO IV, we find that seven country-by-country models, out of 137, and the two-stage post-LASSO IV model at the panel-level satisfy the second criterion by yielding the first-stage F-statistics greater than 10. Note that our estimates from the traditional IV models are unstable at the country-level due to the limited number of monthly observations. Therefore, we are unable to replicate the qualitative findings in Hausman and Kellogg [9] and Davis and Muehlegger [10] for EU data at the country level. However, our findings for three out of 23 EU countries, and our panel-level results, suggest that other countries' weather variables have an indirect relationship with residential natural gas consumption that supports the authors' suggestions. Moreover, Kang et al. [12] argue that the LASSO-selected instrumental variables might be an invalid instruments while they are relatively strong. While we find that the selected instrumental variables from the two-stage post-LASSO IV model are invalid for seven out of 23 countries, our results for first-stage relevance are also not particularly strong at the individual country level, perhaps again because of the limited data on individual countries. We find by contrast that the sets of instrumental variables in IV Model 1 and IV Model 2, which contain the sum of all other countries HDD and CDDs, are valid instrumental variables for these countries.

Our preferred estimates for natural gas price elasticities are obtained from the best models that are selected specific to each country based on the three criteria. Note that, if none of these models satisfy the second criterion, we still select the model that provides the highest first-stage F-statistics to estimate the natural gas price elasticity. Specifically, our study does not suggest reliable estimates for Austria, Latvia, and Slovenia since none of the models satisfy either the first or the third criteria for these countries. Moreover, the results for the remaining 20 EU countries imply that the preferred residential natural gas price elasticities are estimated in the range of $[-0.98, -0.09]$. In addition, for Poland, Portugal, and Slovakia, the best models satisfy all three criteria together implying that the natural gas price elasticities for these countries are more reliable. Specifically, these price elasticities are found to be -0.36, -0.45, and -0.68 for these countries, respectively, showing that natural gas demand is more elastic for Slovakian households than that of Poland and Portugal. In addition, it shows that Poland and Portugal households are in higher risks of getting affected by a possible natural gas supply disruption compared to Slovakian households. One possible solution may be to invest more in renewable energy technologies to increase the available substitutes for residential natural gas. It is worth noting that from 2013 to 2018, Slovakian households' renewable energy consumption had a 12% increase for residential purposes, whereas those of Portugal and Poland had 3% and 9% reductions respectively. However, the transition from natural gas to renewable energy can be challenging with respect to time and money for householders. To this end, governments of the countries where natural gas is more inelastic should be more prepared for supply disruptions and price shocks in order to minimize the impact on their householders' welfare. Increasing supply diversity by investing in new transportation infrastructures in the EU natural gas market, which currently relies primarily on Russian gas, can be a solution for reduce these risks. To this end, in a companion paper we use our preferred country-level natural gas demand elasticity estimates in a highly detailed EU natural gas supply chain model in order to investigate options for market supply diversity through transportation infrastructure [31].

In addition to country-level estimates, panel-level results indicate that the selected set of optimal instrumental variables by two-stage post-LASSO IV have the best relationship with residential natural gas price. According to these estimates, the panel-level natural gas price elasticity for the 23 EU countries in our sample as a whole is -0.32 which implies that residential natural gas is a relatively inelastic good for EU households. An application for future work is to better understand how different LASSO-selected instruments in a panel may be better or worse instruments for different panel units. In our application at the panel-level, the selected set of instrumental variables in two-stage post-LASSO IV and post-LASSO IV models are common to all 23 EU countries simultaneously in a given specification. However, each of these common variables may not have the same effect for all countries. For example, one instrumental variable may be strong for a specific country, while the same variable may not be sufficiently strong for another country. To this end, the application of LASSO methodologies in systems of equations needs more investigation.

CHAPTER 3

OPTIMIZING INVESTMENT AND TRANSPORTATION DECISIONS FOR THE EUROPEAN NATURAL GAS SUPPLY CHAIN

Submitted to *Applied Energy.*

Merve Olmez Turan[1,2], Tulay Flamand[3], Benjamin Gilbert[3]

3.1 Abstract

We address the European natural gas supply chain with several tiers, including producers, mid-streamers, and consumers, where natural gas and Liquefied Natural Gas (LNG) could be traded via long-term contracts or spot markets. This network problem is formulated as a non-linear mixed-integer programming model which prescribes the optimal production and export decisions of producers, import and storage decisions of mid-streamers, and infrastructure investment decisions of European Union (EU) countries with respect to new pipelines and LNG regasification terminals that maximize the total social welfare in the EU over a five-year horizon. We conduct several case studies to examine this network under different conditions. We first compare the actual and optimal decisions to provide insights. Then, we examine the effect of infrastructure decisions on social welfare. Our results reveal that new infrastructure investments increase the total social welfare by nearly three billion. In addition, we examine its sensitivity to the exclusion of Russian gas supply from the market with and without the infrastructure decisions. Results suggest that if Russian gas supply is excluded from the market, than the social welfare and cumulative natural gas consumption of 26 EU countries decrease by 10% and 15%, respectively and that considering infrastructure investments on LNG terminals and pipelines would reduce supply risk of consumer countries.

3.2 Introduction

Natural gas plays a significant role throughout the world as well as Europe. Specifically, Russia's invasion of Ukraine in 2022 has led concerns about EU natural gas supply [2], which makes it necessary to seek and perform appropriate energy strategies. The International Gas Union (IGU) explains the critical role of natural gas with three main reasons: (i) it is cheaper comparing to other fossil-fuel energy sources, (ii) it is more convenient in terms of its infrastructure and delivery, and (iii) it is a sustainable resource that can mitigate climate change and reduce pollution, since it has 50% less CO_2 emissions than those of

[1]Graduate student, Colorado School of Mines
[2]Primary researcher and author
[3]Assistant Professor, Colorado School of Mines

coal and 30% less emissions than those of oil. Due to these factors, many countries aim to improve their natural gas production technologies such as using unconventional production methods, and/or to expand their infrastructure network elements (e.g., storage, pipelines and LNG terminals) in the natural gas supply chain.

According to the IGU, in 2018, nearly $360 billion were invested in natural gas infrastructure across the world [32], since domestic natural gas production has not met the total consumption during the last decade. In 2019, the 10% of the total natural gas demand for 27 EU countries (EU-27) are met by the European natural gas production [1]. In addition, the annual gap between natural gas consumption and production of these countries increased from 304 billion cubic meters (bcm) to 409 bcm between 2014 and 2020 [1]. Due to this reason, these countries imported nearly half of its total gas consumption from several other countries. Particularly, in 2019, 35% of natural gas demand of these European countries was supplied by Russian Federation. Norway, Algeria, Qatar, and the United States (US) also supplied 20%, 5.5%, 5.1%, and 4.3% of the total demand, respectively [1]. Russia has been the leading supplier of the natural gas market in Europe. Although the US increased its LNG export volume from 3.9 bcm to 22.8 bcm in 2020, Russia remains its position as the primary gas supplier for both LNG and pipeline markets in Europe with the total natural gas volume of 184.7 bcm [1].

Most pipelines in the EU are currently operating above 80% of their monthly capacity peaks [33]. Furthermore, 75% of the additional LNG import between September 2018 and March 2019 was used to satisfy the demand of the countries in the northwest Europe [33]. However, some regions in the central and southeast Europe still encounter difficulties in accessing alternative gas supplies due to their limited pipeline connections to supplier countries. In addition, there are LNG regasification terminals in only about half of the European countries (16 out of 29) and most of them are located in the southwest of Europe. In order to increase accessibility of the natural gas for European countries, particularly, by those in the central and southeast Europe, the European Network of Transmission System Operators for Gas (ENTSO-G) is planning to open new pipelines and LNG terminals.

For the countries which import natural gas, energy security is an important concern. During the mid-winter of 2009, since Ukraine blocked Russian natural gas to be transported to Europe through its pipelines due to a price dispute of natural gas between them, European countries had a serious supply shortage problem [34]. Moreover, Russia's current invasion of Ukraine has also increased these concerns about the EU natural gas reliance on Russia [2]. Although during the invasion, the gas flow from Russia to EU countries has not interrupted, EU countries have been significantly impacted due to an increase in natural gas price. In addition, there are three subjects that can affect the EU's gas security: (i) the capacity of gas flow within the EU should be evaluated; (ii) the demand pattern in the future should be forecasted; and (iii) the gas usage in reducing the carbon footprints in the energy systems should be studied [35]. In order to ensure that the energy supply is secure, sustainable, competitive, and affordable for European consumers, in 2015, the Energy Union policy was announced [36] which covers a fundamental transformation for the energy system of Europe. Specifically, it aims to ensure energy security, solidarity, and trust; a fully integrated European energy market; energy efficiency; decarbonization; and the continuity of research, innovation, and competitiveness [36]. Despite all these efforts that intends to reduce the energy dependency, Russia is still expected to remain its position as a main natural gas supplier in the market until 2040, with a 30% market share [33].

In this paper, we seek to answer the following research questions: (i) What are the optimal production and export decisions of producers, import and storage decisions of mid-streamers, and infrastructure decisions of the regulator that jointly maximize the social welfare in the European natural gas market? (ii) What is the effect of these investment decisions on the social welfare? (iii) How much do these investment decisions change the market share of producers and market structure? (iv) How the social welfare changes if Russia was not considered as a supplier in the EU natural gas market? In order to answer these questions, we develop a non-linear mixed-integer optimization model which prescribes the optimal infrastructure decisions of countries as well as their optimal trade and storage volumes and market price. The proposed model considers the European natural gas market for 26 EU countries as consumers and 12 countries as producers during the time horizon between 2020 and 2025.

The remainder of this paper is organized as follows: Section 2 describes the literature review. Section 3 introduces a formal problem statement, model assumptions, notation and it proposes an optimization model. Section 4 describes our computational study along with the description of data, computational experiment and discussion of the results. Section 5 concludes the paper and discusses the future work.

3.3 Literature Review

Many researchers propose various models and use different methodologies in order to examine the natural gas supply chain. Almost all models predict traded natural gas volumes among countries under market-clearing conditions. Egging and Gabriel [37] investigate the effect of market power on market equilibrium under four scenarios, namely, perfect competition, oligopoly producer, three extra pipelines connected to the United Kingdom, no storage. Egging et al. [38] also examine the scenario where there is a disruption of supplies coming from Russia through Ukraine and Algeria. The authors suggest that pursuing sufficient import pipeline capacity and supplying from diverse suppliers are critical for energy security of countries that are mostly dependent on the European pipeline market. Allevi et al. [39] consider the effect of long-term contract negotiations for the Italian natural gas market, when spot market prices are lower than long-term contract prices. Their findings reveal that mid-streamers would end up with a significant profit loss, if long-term contract prices are not changed by negotiations, while in the case of a price change, the availability of natural gas is reduced that also causes a reduction in the social welfare [39].

To analyze such decisions considered in the aforementioned studies, several models are used, namely, mixed-complementary models [38, 40–42], general equilibrium models [43–45], stochastic models [46, 47], dynamic programming [48], mixed-integer non-linear optimization models [39] and mixed-integer linear optimization models [49, 50]. In the literature, these models are generally used to achieve two different objectives; profit maximization and social welfare maximization. The objective of profit maximization is mostly used in mixed-complementary models for individual actors under different market structures [40–42, 51], while the objective of social welfare maximization is used in the mixed-integer non-linear models [39] and general equilibrium natural gas models [43–45] under the assumption of the perfect competition market.

In the literature, there are several studies that consider infrastructure decisions endogenously for the natural gas market, such as RAMONA, European Gas Market Model (EGMM), and Global Gas Model (GGM). Fodstad et al. [52] proposes RAMONA, which is a multi-stage stochastic optimization model that prescribes the optimal infrastructure decisions for a perfect competition market in order to maximize the social welfare between 2010 and 2050 with a five-year resolution. Kiss et al. [53] propose EGMM which is a multi-market equilibrium model that examines infrastructure decisions individually for 35 European countries to maximize the social welfare under perfect competition. GGM is a mixed-complementary model that provides infrastructure investment decisions in order to maximize the total profit for producers, traders, transmission system operators, and storage operators [45].

Our work is most related in this regard to Lochran [49] that proposes a mixed-integer linear programming model, GNOME. The proposed model prescribes optimal investment decisions regarding pipelines and LNG regasification terminals in Europe in order to minimize the total cost. The author considers long-term contracts as well as storage decisions in GNOME. However, the natural gas demand, which has a dynamic relationship with supply cost, is considered as single deterministic parameter for each country in GNOME. By contrast, we examine natural gas for three sectors separately, namely, household, industry, and power and these sectors for each country have different demand functions in order to prescribe market-equilibrium points, endogenously.

In this paper, we examine the European natural gas supply chain as a network model. Our contributions are as follows: (i) To the best of our knowledge, this is the first study which proposes a mixed-integer non-linear programming model that jointly considers trade and storage decisions of all supply chain components (i.e., producers, mid-streamers, consumers) via long-term contracts or spot markets as well as infrastructure decisions of the regulator regarding opening new pipelines and LNG regasification terminals that maximizes the consumer welfare for 26 EU countries from the regulator perspective; (ii) it is the first study that investigates the effect of infrastructure investments on market structures at the extraction stage of European natural gas market; and (iii) that examines how the social welfare changes if Russia was not considered as a supplier in the EU natural gas market.

3.4 The European Gas Supply Chain Model

In this section, we provide our formal problem statement and a mixed-integer non-linear programming model for the European natural gas supply chain.

3.4.1 Problem Statement

In this paper, we examine the European natural gas supply chain as a network model that includes producer and consumer nodes. Producer countries are the Russian Federation (RU), Norway (NO), Libya (LY), Algeria (DZ), Qatar (QA), Egypt (EG), Nigeria (NG), Peru (PE), the United States (US), the United Kingdom (UK), Netherland (NL), and Trinidad and Tobago (TT), while the consumers are the UK and 26 EU countries (except for Malta and Cyprus), resulting in a total of 38 countries.

A sample representation of the natural gas supply chain is shown in Figure 3.1 as a network with one producer and two consumer nodes. Each producer extracts the natural gas from its production well. The extracted gas may either be sent directly to consumers via pipelines, as highlighted with red arrows, or may be first processed in its LNG liquefaction terminal to be converted into LNG, and be shipped to LNG regasification terminals of consumers, as highlighted in green arrows, to be regasified. We consider that supply capacity of producers as well as capacity of LNG liquefaction terminals are fixed capacities without a change over time. We consider each consumer with a single mid-streamer. Mid-streamers connect the market and end-users by trading the natural gas. Producers may sell the natural gas to mid-streamers via long-term contracts, or spot markets. Note that, long-term contracts enforce mid-streamers to purchase a specified minimum amount of gas in a given time period. We consider each consumer with at most one spot market. Mid-streamers may purchase and/or sell the natural gas from (to) either the spot market of its country or those of other countries via pipelines. Each mid-streamer may have storage that allows them to import and store the natural gas when its price is cheaper. End-user of each consumer represents the final destination of the natural gas, where it is consumed in three sectors, namely, household, power generation and industry. All sectors are considered with a linear inverse demand function. The gas is transported between consumers via pipelines.

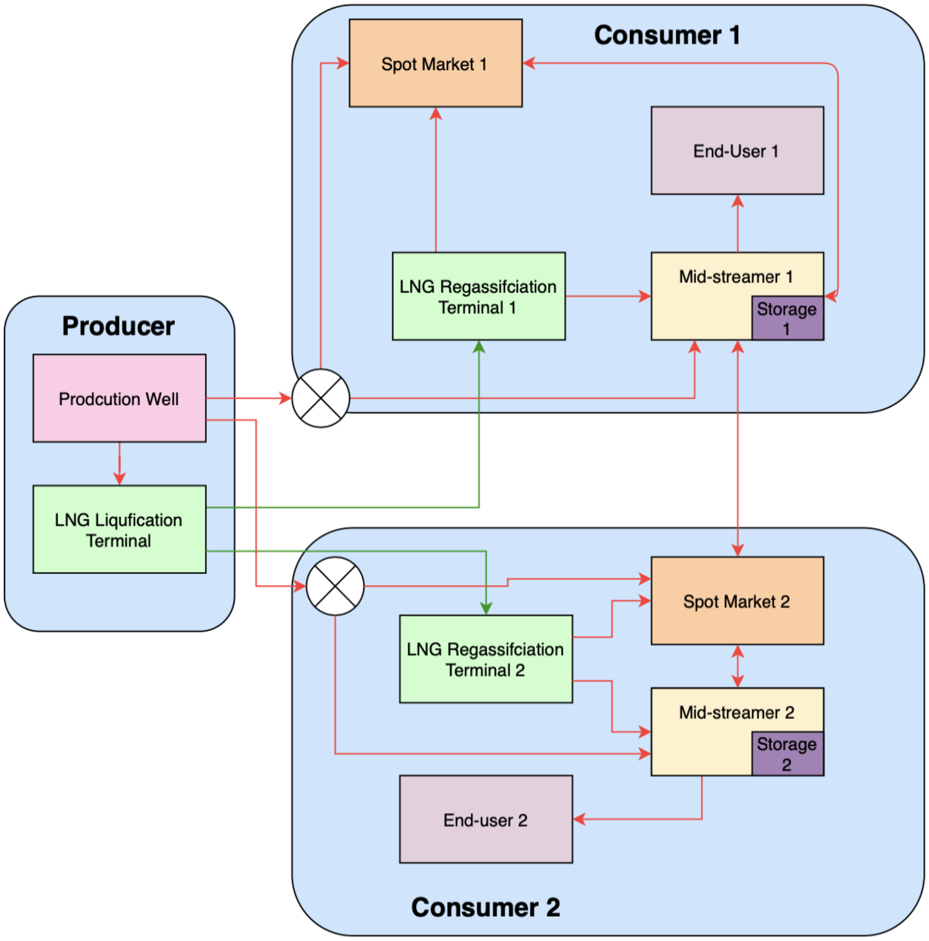


Figure 3.1 The representation of the natural gas supply chain

In the next section, we examine this problem for market-clearing conditions by proposing a mixed-integer non-linear optimization model in order to maximize the social welfare.

3.4.2 Mathematical Formulation

In this section, we develop a mixed-integer non-linear programming model that prescribes the optimal trade, storage decisions of producers, mid-streamers and consumers in the European natural gas supply chain via long-term contracts or spot markets, as well as infrastructure decisions of opening new pipelines and LNG regasification terminals. This model aims to maximize discounted social welfare for the time horizon from 2020 to 2025 with a monthly resolution. In addition, the infrastructure decisions are considered to be made annually.
Our notation and the proposed mathematical formulation, European Gas Infrastructure Model (**EGIM**), are introduced below. Consider the sets and parameters as follows:

Sets	
$\mathcal{P}$	Producer nodes
$\mathcal{C}$	Consumer nodes
$\mathcal{N}_c$	Spot markets, mid-streamers, and end-users in consumer node c
$\mathcal{M}$	Mid-streamers
$\mathcal{S}$	Spot markets
$\mathcal{E}$	End-users
$\mathcal{Z}$	Sectors of end-users (Household=1, Industry=2, Power=3)
$\mathcal{R}$	Transportation methods (Pipeline=1, LNG=2)
$\mathcal{L}$	Potential LNG regasification terminals
$\mathcal{K}$	Potential pipelines
$\mathcal{F}_{\mathrm{r}}$	Producer-consumer pairs (p, c) that has connection from producer p to consumer c via transportation method r
$\mathcal{F}^{\mathrm{K}}$	Triplets (p, c, k) that represents producer-consumer pairs (p, c) that may get a capacity expansion by opening new pipelines with project k
$\mathcal{J}$	Consumer pairs (c_1, c_2) that already have or potentially would have pipeline connections
$\mathcal{J}^{\mathrm{K}}$	Triplets (c_1, c_2, k) that represents the consumer pairs (c_1, c_2) that may get a capacity expansion by opening new pipelines with project k
$\mathcal{A}$	Triplets (i, j, r) that represents a connection between i and j via transportation method r according to one of these following conditions: $(i, j, r) \in \mathcal{A}^{LT}$, $i \in \mathcal{P}, j \in \mathcal{S}, r \in \mathcal{R}$, $i \in \mathcal{M}, j \in \mathcal{S}, r = 1$, $i \in \mathcal{M}, j \in \mathcal{E}, r = 1$, $i \in \mathcal{S}, j \in \mathcal{M}, r = 1$
$\mathcal{A}^{\mathrm{LT}}$	Triplets (i, j, r) that represents a connection between producer p and mid-streamer m using a long-term contract via transportation method r
$\mathcal{Y}$	Quadruplets (o, f, i, j) that represents a connection between origin producer o to final (destination) consumer f that could be accessed by passing an intermediate connection (i, j) via pipeline
$\mathcal{Y}^{\mathrm{LT}}$	Origin producer o and final consumer f pairs, (o, f), that does not have a direct connection, i.e., it is necessary to use one or more intermediate connections. Here, it is considered that all connections are via pipeline and the purchase between o and f occurred via a long-term contract
$\mathcal{Y}^{\mathrm{IC}}$	Pairs (i, j) that represents an intermediate connection between origin producer o and final consumer f via pipeline

Subsets Sets	
$\mathcal{P}^{\mathrm{LNG}} \subset \mathcal{P}$	Set of producer nodes with an LNG liquefaction terminal
$\mathcal{C}^{\mathrm{LNG}} \subseteq \mathcal{C}$	Consumer nodes that already or potentially have an LNG regasification terminal
$\mathcal{N}_{\mathrm{c}}^{\mathrm{S}} \subseteq \mathcal{N}_c$	Spot markets in consumer node c
$\mathcal{N}_{\mathrm{c}}^{\mathrm{M}} \subseteq \mathcal{N}_c$	Mid-streamers in consumer node c
$\mathcal{N}_{\mathrm{c}}^{\mathrm{E}} \subseteq \mathcal{N}_c$	End-users in consumer node c

In addition, consider the model parameters below:

Cost Parameters

c_p^{PR}	Marginal production cost of producer p	[€ / mcm]
c_{ijr}	Marginal transportation cost of using connection (i,j) via transportation method r	[€ / mcm]
c_m^{INV}	Monthly marginal unit holding cost of mid-streamer m	[€ / mcm]
c_m^{INV+}	Monthly marginal unit injection cost of natural gas for the storage of mid-streamer m	[€ / mcm]
c_m^{INV-}	Monthly marginal unit withdrawal cost of natural gas for the storage of mid-streamer m	[€ / mcm]
c_k^F	Fixed investment cost of opening a pipeline or an LNG regasification terminal with project k	[€]

Capacity Parameters

C_p^R	Reserve capacity of producer p	[mcm]
C_p^{PR}	Monthly production capacity of producer p	[mcm]
C_p^{LQ}	Monthly LNG liquefaction capacity of producer p	[mcm]
C_c^{RQ}	Monthly capacity of the current LNG regasification terminals for consumer c	[mcm]
C_{cl}^{PRQ}	Monthly capacity of the potential LNG regasification terminal project l for consumer c	[mcm]
C_{ij}^{PL}	Monthly capacity of the current pipelines on connection (i,j)	[mcm]
C_{ijk}^{PPL}	Monthly capacity of the potential pipeline project k on connection (i,j)	[mcm]

Storage Parameters

C_m^{INV}	Total capacity of storage for mid-streamer m	[mcm]
C_m^{INV+}	Monthly capacity of the injection amount for the storage of mid-streamer m	[mcm]
C_m^{INV-}	Monthly capacity of the withdrawal amount for the storage of mid-streamer m	[mcm]
II_m^0	Initial storage amount for mid-streamer m	[mcm]

Other Parameters

α_p	Rate of LNG liquefaction loss for producer p	[%]
β_c	Rate of LNG regasification loss for consumer c	[%]
g_t	Available budget of year t for opening potential pipelines and LNG terminals	[€]
l_{pmrt}	Minimum purchase volume for the long-term contract between producer p and mid-streamer m via transportation method r during year t	[mcm]
δ_t	Discount rate of month t	[%]
$a_{e,z,t}$	Constant parameter in demand function of end-user e for sector s in month t	[%]
$b_{e,z,t}$	Consumption coefficient in the demand function of end-user e for sector s in month t	[%]

The decision variables of Model **EGIM** are introduced below:

Continuous Variables

X_{ijrd}	Amount of natural gas transported using connection (i,j,r) in month d	[mcm]

γ_{ofijd}	Amount of natural gas transported using connection (o, f, i, j) that represents the shipment from producer o to consumer f by passing an intermediate connection (i, j) via pipeline in month d	[mcm]
I_{md}	Amount of natural gas available in the storage of mid-streamer m at the end of month d	[mcm]
I^+_{md}	Amount of natural gas injected to the storage of mid-streamer m, at the end of month d	[mcm]
I^-_{md}	Amount of natural gas withdrawn from the storage of mid-streamer m, at the end of month d	[mcm]
Q_{ezd}	Amount of natural gas consumed by sector z of end-user e in month d	[mcm]
Binary Variables		
z_{kt}	$z_{kt} = 1$ if and only if potential pipeline or LNG regasification terminal project k is opened at the beginning of year t	$\{0, 1\}$
y_{md}	$y_{md} = 1$ if and only if there is an injection to the storage of mid-streamer m at the end of month d	$\{0, 1\}$

Model **EGIM** is then formulated as the following mixed-integer non-linear programming model:

$$\text{Maximize} \sum_{e\in\mathcal{E}}\sum_{z\in\mathcal{Z}}\sum_{d\in\mathcal{D}}(\delta_t \cdot a_{ezd} \cdot Q_{ezd} + \delta_t \cdot 0.5 \cdot b_{ezd} \cdot Q^2_{ezd}) \tag{3.1a}$$

$$- \sum_{(p,j,r)\in\mathcal{A}:p\in\mathcal{P}}\sum_{d\in\mathcal{D}} \delta_t \cdot c^{PR}_p X_{pjrd} - \sum_{(p,m,p,j)\in\mathcal{Y}:p\in\mathcal{P}}\sum_{d\in\mathcal{D}} \delta_t \cdot c^{PR}_p \gamma_{pmpjd} \tag{3.1b}$$

$$- \sum_{(i,j,r)\in\mathcal{A}}\sum_{d\in\mathcal{D}} \delta_t \cdot c_{ijr} X_{ijrd} - \sum_{(p,m,i,j)\in\mathcal{Y}}\sum_{d\in\mathcal{D}} \delta_t \cdot c_{ij1}\gamma_{pmijd} \tag{3.1c}$$

$$- \sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \delta_t \cdot c^{INV+}_m I^+_{md} - \sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \delta_t \cdot c^{INV-}_m I^-_{md}$$

$$- \sum_{m\in\mathcal{M}}\sum_{d\in\mathcal{D}} \delta_t \cdot c^{INV}_m I_{md} \tag{3.1d}$$

Subject to:

$$\sum_{(p,j,r)\in\mathcal{A}}\sum_{d\in\mathcal{D}} X_{pjrd} + \sum_{(p,m,p,c)\in\mathcal{Y}}\sum_{d\in\mathcal{D}} \gamma_{pmpcd} \leq C^R_p, \qquad \forall p \in \mathcal{P} \tag{3.2a}$$

$$\sum_{(p,j,r)\in\mathcal{A}} X_{pjrd} + \sum_{(p,m,p,c)\in\mathcal{Y}} \gamma_{pmpcd} \leq C^{PR}_p, \qquad \forall p \in \mathcal{P}, d \in \mathcal{D} \tag{3.2b}$$

$$\sum_{(p,j,2)\in\mathcal{A}} X_{pj2d} \leq C^{LQ}_p, \qquad \forall p \in \mathcal{P}^{LNG}, d \in \mathcal{D} \tag{3.2c}$$

$$\sum_{p\in\mathcal{P}:(p,c)\in\mathcal{F}_2}\sum_{j\in\mathcal{N}^M_c\cup\mathcal{N}^S_c}(1-\alpha_p)X_{pj2d} \leq C^{RQ}_c + \sum_{k\in\mathcal{L}}\sum_{t'=1}^{t} C^{PRQ}_{ck} z_{kt'},$$
$$\forall c \in \mathcal{C}^{LNG}, t \in \mathcal{T}, d \in \mathcal{D}_t \tag{3.2d}$$

$$\sum_{(p,m,c,c')\in\mathcal{Y}} \gamma_{pmcc'd} + \sum_{(p,m,c,m)\in\mathcal{Y}:m\in\mathcal{N}^M_{c'}} \gamma_{pmcmd} + \sum_{j\in\mathcal{N}^M_c}\sum_{j'\in\mathcal{N}^S_{c'}} X_{jj'1d}$$

$$+\sum_{j\in\mathcal{N}_c^S}\sum_{j'\in\mathcal{N}_{c'}^M} X_{jj'1d} \le C_{cc'}^{PL} + \sum_{(c,c',k)\in\mathcal{J}^K}\sum_{t'=1}^{t} C_{cc'k}^{PPL} z_{kt'},$$
$$\forall (c,c') \in \mathcal{J}, t\in\mathcal{T}, d\in\mathcal{D}_t \tag{3.2e}$$

$$\sum_{j\in\mathcal{N}_c^M\cup\mathcal{N}_c^S} X_{pj1d} + \sum_{(p,m,p,c)\in\mathcal{Y}} \gamma_{pmpcd} \le C_{pc}^{PL} + \sum_{(p,c,k)\in\mathcal{F}^K}\sum_{t'=1}^{t} C_{pck}^{PPL} z_{kt'},$$
$$\forall (p,c)\in\mathcal{F}_1, t\in\mathcal{T}, d\in\mathcal{D}_t \tag{3.2f}$$

$$\sum_{t\in\mathcal{T}} z_{kt} \le 1, \qquad \forall k\in\mathcal{L}\cup\mathcal{K} \tag{3.2g}$$

$$I_{md}^{+} \ge \sum_{(i,m,2)\in\mathcal{A}} (1-\alpha_i)(1-\beta_m)X_{im2d} + \sum_{(i,m,1)\in\mathcal{A}} X_{im1d}$$
$$- \sum_{(m,j,1)\in\mathcal{A}} X_{mj1d} + \sum_{(p,m,k,m)\in\mathcal{Y}} \gamma_{pmkmd},$$
$$\forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2h}$$

$$I_{md}^{-} \ge \sum_{(m,j,1)\in\mathcal{A}} X_{mj1d} - \sum_{(i,m,2)\in\mathcal{A}} (1-\alpha_i)(1-\beta_m)X_{im2d}$$
$$- \sum_{(i,m,1)\in\mathcal{A}} X_{im1d} - \sum_{(p,m,k,m)\in\mathcal{Y}} \gamma_{pmkmd},$$
$$\forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2i}$$

$$I_{md} = I_{m(d-1)} + I_{md}^{+} - I_{md}^{-}, \qquad \forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2j}$$

$$I_{md} \le C_m^{INV}, \qquad \forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2k}$$

$$I_{md}^{+} \le C_m^{INV+} y_{md}, \qquad \forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2l}$$

$$I_{md}^{-} \le C_m^{INV-}(1-y_{md}), \qquad \forall m\in\mathcal{M}, d\in\mathcal{D} \tag{3.2m}$$

$$I_{m0} = II_m^0, \qquad \forall m\in\mathcal{M} \tag{3.2n}$$

$$\sum_{(i,s,1)\in\mathcal{A}} X_{isr1} + \sum_{(i,s,2)\in\mathcal{A}} (1-\alpha_i)(1-\beta_s)X_{is2d} = \sum_{(s,j,1)\in\mathcal{A}} X_{sj1d},$$
$$\forall s\in\mathcal{S}, d\in\mathcal{D} \tag{3.2o}$$

$$\sum_{(m,e,1)\in\mathcal{A}} X_{me1d} = \sum_{z\in\mathcal{Z}} Q_{ezd}, \qquad \forall e\in\mathcal{E}, d\in\mathcal{D} \tag{3.2p}$$

$$\sum_{d\in\mathcal{D}_t} X_{pmrd} \ge l_{pmrt}, \qquad \forall (p,m,r)\in\mathcal{A}^{LT}, t\in\mathcal{T} \tag{3.2q}$$

$$\sum_{(p',m,p',j)\in\mathcal{Y}:p'=p}\sum_{d\in\mathcal{D}_t} \gamma_{p'mp'jd} \ge l_{pm1t}, \qquad \forall (p,m)\in\mathcal{Y}^{LT}, t\in\mathcal{T} \tag{3.2r}$$

$$\gamma_{pmijd} = \sum_{(p,m,j,k)\in\mathcal{Y}} \gamma_{pmjkd}, \qquad \forall (p,m,i,j)\in\mathcal{Y}, d\in\mathcal{D} \tag{3.2s}$$

$$\sum_{k\in\mathcal{L}\cup\mathcal{K}} c_k^F z_{kt} \le t\cdot g_t - \sum_{k\in\mathcal{L}\cup\mathcal{K}}\sum_{t'=1}^{t} c_k^F z_{kt'}, \qquad \forall t\in\mathcal{T} \tag{3.2t}$$

$$x, \gamma, I, I^{+}, I^{-}, q \ge 0, z, y \text{ binary} \tag{3.2u}$$

The objective function (3.1a)-(3.1d) maximizes the discounted end-user social welfare. Let $P_{ezd}(q)$ represent the inverse demand function for the demand of sector z of end-user e in month d which is equal to $a_{ezd} + b_{ezd}\cdot Q_{ezd}$. Then the first term of the objective function, (3.1a), represents the total consumer surplus which is the summation of the $\int_0^{Q_{ezd}} P_{ezd}(q)dq$ over all $e \in \mathcal{E}, z \in \mathcal{Z}, d \in \mathcal{D}$. This can also be written as $a_{ezd} \cdot Q_{ezd} + 0.5 \cdot b_{ezd} \cdot Q^2_{ezd}$, where a and b are the demand function parameters to be estimated. Terms (3.1b), (3.1c) and (3.1d) of the objective function represent the total production cost of producers, the total transportation cost of the supply chain, and the total storage injection, withdrawal, holding cost of mid-streamers, respectively. Constraint (3.2a) enforces that the total production of producer p during the entire time horizon cannot exceed its total reserve capacity. Constraint (3.2b) ensures that monthly production of producer p cannot exceed its monthly production capacity. Note that, given Constraint (3.2b), Constraint (3.2a) would be redundant, particularly, if the problem is considered for a time horizon which is smaller than the reserve lifetime. Constraint (3.2c) ensures that the total amount of LNG transported from a producer p to consumers in a month, does not exceed its monthly LNG liquefaction capacity. Constraint (3.2d) ensures that the total amount of LNG to be regasified, received by the mid-streamer and spot markets of consumer c from all producers in a month cannot exceed its monthly LNG regasification capacity, including the monthly capacity of the current LNG regasification terminals and the capacity of potential LNG regasification terminals, if opened. Here, the total amount of LNG received by consumer c is calculated by taking into consideration the liquefaction losses of producers. Constraint (3.2e) ensures that the total amount of natural gas transported via pipelines that connect each pair of consumers (c, c') cannot exceed the the monthly pipeline capacity, including the monthly capacity of the current pipelines and the capacity of potential pipelines, if opened. Constraint (3.2f) ensures the same condition for pipelines that connect every producer-consumer pair (p, c). Constraint (3.2g) ensures that each potential pipeline or LNG regasification terminal may be opened at most once. Constraints (3.2h) and (3.2i) jointly calculate the monthly injection or withdrawal storage amount of each mid-streamer m, whichever occurs. Particularly, injection occurs if the received amount of gas for mid-streamer m is larger than the transported amount. Otherwise, some amount of gas is withdrawn from the storage. Constraint (3.2j) balances the ending storage of the current month by considering the starting storage of the month (i.e., ending storage of the previous month) and injected or withdrawn amounts, whichever occurs. Constraints (3.2k)-(3.2m) ensure that the monthly storage, injection and withdrawn volumes do not exceed their corresponding capacities, respectively. Note that the auxiliary variable y is introduced here in order to enforce that either injection or withdrawal would occur for the storage of a mid-streamer, but not simultaneously. Since in the objective function withdrawal and injection costs are minimized, one would expect that, one of them would occur naturally without a need of the auxiliary variable. However, in our data, the withdrawal cost, c_m^{INV-}, is zero. For the instances where injection and withdrawal costs are non-zero, this auxiliary variable is not needed. Constraint (3.2n) assigns the initial storage volumes to mid-streamers. Constraints (3.2o) and (3.2p) ensure flow-balance for each spot market and each end-user, respectively. Particularly, total natural gas amount received by a spot market should be equal to the total amount of gas transported from it. Similarly, the total amount of natural gas purchased by an end-user from its mid-streamer should be equal to all sectors' demand of this end-user. Constraints (3.2q) and (3.2r) ensure that each mid-streamer purchases from each producer via a long-term contract must be at least as much as the specified minimum annual contract volume. Particularly, Constraint (3.2q) ensures this condition for each pair of producer and mid-streamer that are directly connected, while, Constraint (3.2r) ensures it for each pair of producer and mid-streamer that do not have a direct connection. Constraint (3.2s) ensures that countries that are intermediate connections between an origin and a destination country must transfer the received volume to the target destination. Constraint (3.2t) ensures that the total investment on new projects does not exceed the annual budget. Note that, annual budget also covers the rollover of unused budget from previous years. Constraint (3.2u) enforces non-negativity and binary restrictions on decision variables.

3.5 Computational Study

In this section, we detail our data acquisition, and conduct a computational study by implementing the proposed model on the obtained data. We provide and discuss our results.

3.5.1 Data Set

In this section, we provide details regarding the data gathered and used as parameters in the Model **EGIM**. The proposed model considers the natural gas market as a network including set of producer and consumer nodes. Specifically, we consider 12 producers; six of them export their production only in the form of LNG (Qatar (QA), the United States (US), Peru (PE), Nigeria (NG), Egypt (EG), and Trinidad and Tobago (TT)), while three of them export their production only via pipelines (Libya (LY), Netherland (NL), and the United Kingdom (UK)), and three of them export the natural gas either via LNG or pipeline (the Russian Federation (RU), Norway (NO), and Algeria (DZ)).

We obtain natural gas reserve volumes of producers and their actual production volumes from BP [54]. As mentioned in Egging et al. [38], precise information for production capacity of producers is not available. To this end, in this paper, we follow the capacity calculation of Egging et al. [38] by taking 1.1 times of the actual production volumes in 2019. This assumption is made since all producers operate with high efficiency by using a large percentage of their capacities. LNG liquefaction capacities of producers are obtained from GIE [55][dataset]. Moreover, capacities of the existing LNG regasification terminals and pipelines are gathered from ENTSO-G [56][dataset]. We also consider 16 potential pipelines and four LNG regasification terminals on which can be invested. These potential projects are obtained from Kotek et al. [57]. Details of these potential projects are given in Table B.1 and Table B.2 in B.

We also consider mid-streamer storage for consumer countries. Storage capacities and available (initial) storage volumes are obtained from GIE [55][dataset]. Moreover, the marginal storage injection and holding costs are obtained from Egging et al. [38]. We consider that there is no cost associated with the storage withdrawal as in Egging et al. [38].

The natural gas production cost function of producers is assumed to be linear. The marginal production cost values are obtained from Egging et al. [38]. We obtain data regarding transportation costs for pipelines from Egging et al. [38]. In particular, we consider €10,000/mcm as the marginal overland pipeline cost, €20,000/mcm as the marginal undersea pipeline cost, and €30,000/mcm as the marginal long-distance pipeline cost. In addition, marginal LNG liquefaction and regasification costs are obtained from Baron et al. [58]. Shipping cost is calculated based on the distance between countries, and distance information is obtained from [59]. The unit distance cost per one thousand sea miles is considered €5,000/mcm, as in Egging et al. [38].

Demand function of an end-user for each sector is modeled as an inverse linear function. Specifically, we use the residential natural gas demand elasticity that are estimated by Olmez Turan et al. [60]. We also consider -0.4 as the demand elasticity of the European natural gas market for the industrial sector, and -0.75 as that for power generation, suggested by Egging and Gabriel [37]. Note that the information regarding the monthly natural gas consumption of each end-user for each sector is not available. Therefore, we consider the relative consumption share of household, industry, and electric sectors as 36.7%, 25.8%, and 31.5%, respectively on the total consumption [61][dataset]. The historical natural gas consumption and price information for each end-user are also obtained from IEA [61][dataset].

Finally, we obtain information regarding 426 long-term natural gas contracts, (i.e., participating countries, contract timeline, contract volumes) from Neumann et al. [62]. Specifically, we consider those that span timelines between 2020 and 2025.

3.5.2 Computational Results

In this section, we conduct several case studies as summarized in Table 3.4: (i) a scenario in which we compare actual (historical) demand with our results considering 2019 as the base year; (ii) the market without infrastructure decisions; (iii) the market with infrastructure decisions; (iv) the market without infrastructure decisions and without the consideration of Russian natural gas supply; and (v) the market with infrastructure decisions and without the consideration of Russian natural gas supply. Based on our model parameters, **EGIM** includes $17,607$ constraints and $29,398$ variables, of which $1,046$ are binary variables. The Model **EGIM** is coded in AMPL and solved using CPLEX 20.1.0.0. All computational runs are made on an Apple Mac Book Pro with Intel Core i5-2600 CPU 1.40 GHz processor and 8 GB of RAM. Moreover, CPLEX provides an optimal solution for all case studies within the range of 2 to 3 minutes.

Table 3.4 Case study descriptions

Case	Infrastructure decisions	Russian gas supply	Period
Case 1	No	Considered	2019
Case 2	No	Considered	2020-2025
Case 3	Yes	Considered	2020-2025
Case 4	No	Not considered	2020-2025
Case 5	Yes	Not considered	2020-2025

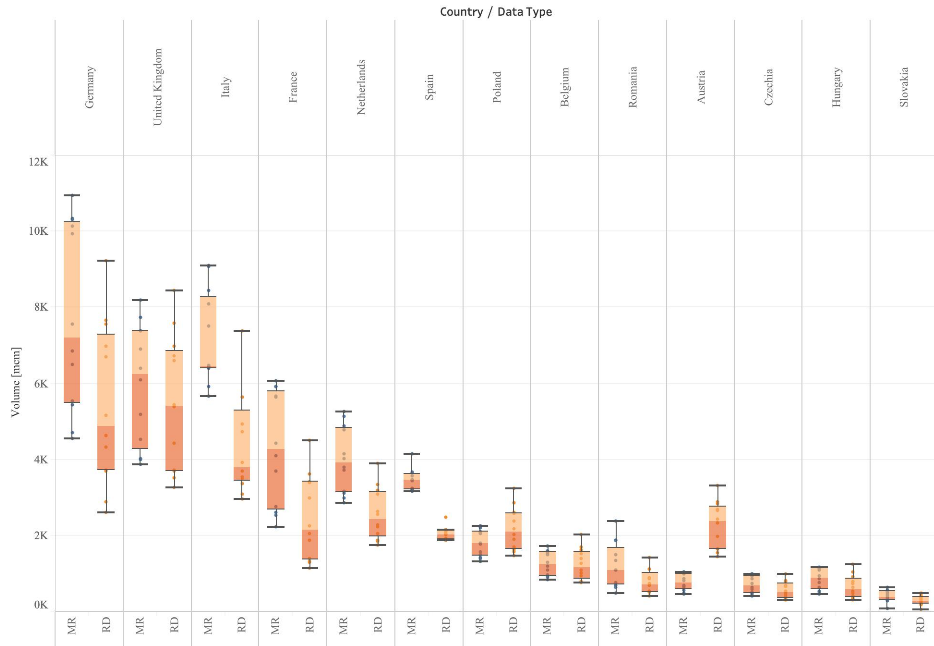
3.5.2.1 Comparison of Actual Consumption in 2019 with Model Results

In this section, in order to compare actual natural gas consumption with optimal results, we run Model **EGIM** for a one-year period of 2019 (Case 1). To this end, we first estimate parameters of inverse demand function and incorporate them into the objective function of the proposed model. Then, by solving the proposed optimization model, we obtain the optimal demand volume of each end-user for each sector. Figure 3.2 compares the actual consumption volumes (represented as RD on x-axis) with the optimal demand volumes obtained by the proposed model (represented as MR on x-axis) for each consumer country. It suggests that the optimal natural gas consumption of Germany, Italy, France, the Netherlands, and Spain are higher than their actual consumption, while the optimal consumption of Austria, Denmark, Lithuania, Finland, and Greece are lower than their actual consumption values in order to maximize the social welfare of 26 EU countries.

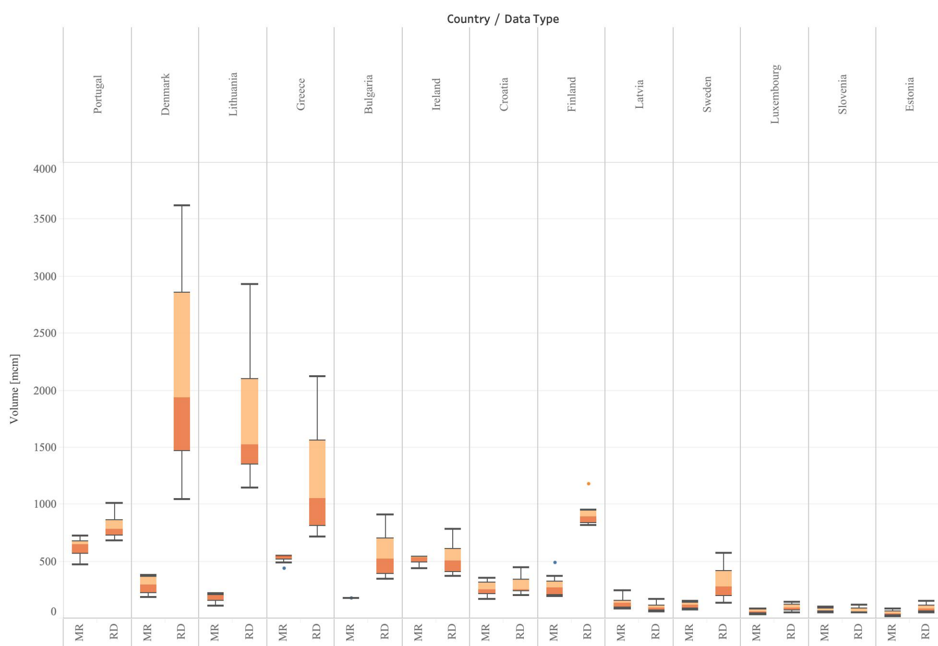
3.5.2.2 Infrastructure Decisions

In this section, we analyze the European natural gas network model using two case studies in order to examine the effect of infrastructure decisions on opening new pipeline and/or LNG regasification terminals. To this end, we consider two cases with and without infrastructure decisions. Particularly, Case 3 considers infrastructure decisions along with trade and storage decisions of all supply chain components, while Case 2 only considers trade and storage decisions without taking into consideration infrastructure decisions. Note that both cases consider a time horizon that spans five years between 2020-2025.

The proposed model prescribes the maximum social welfare of 135,207 million for Case 2, where only current pipeline and LNG capacities are considered without the inclusion of infrastructure decisions (i.e., no opportunity of expanding current capacities exists). In contrast, the maximum social welfare prescribed



(a)



(b)

Figure 3.2 Comparison of actual and optimal consumption for consumer countries for Case 3 is 138,309 million, where the infrastructure decisions are also taken into account with 2,000 million € /year budget. We also calculate the value of country-level Herfindahl-Hirschman Index (HHI) for extraction stage based on the optimal decisions of Cases 2-3 prescribed by the proposed model. In particular, for Case 2 (without infrastructure decisions), HHI is 1,700, while that for Case 3 (with infrastructure decisions) is 1,719. This reveals that expanding current pipeline and LNG capacities for countries by making infrastructure decisions affects the market shares of producers. In other words, the market becomes less competitive if new pipeline and LNG regasification terminals can be opened. Note that for both case studies, the optimal solution suggests that the Russian Federation and Norway have the largest market shares. Optimal investment decisions reduce the market share of Norway, while it increases the market share of Qatar.

Comparing the optimal solutions of Cases 2-3 shows that the total available natural gas is increased from 2,622,919 mcm to 2,738,279 mcm if potential infrastructure investments are considered. In particular, the percentage change of total natural gas consumption of each consumer country is shown in a map in Figure 3.3. This percentage is calculated by first taking difference between monthly natural gas consumption with and without the infrastructure decisions by monthly consumption without infrastructure decisions, then averaging all these percentage differences across five year for each consumer country. Note that color coding is given with respect to the percentage change in consumption, as detailed in Figure 3.3, and countries highlighted in grey are out of scope for this paper. It shows that Bulgaria's optimal natural gas consumption is five times higher in Case 3 (with infrastructure decisions) than the one in Case 2 (without infrastructure decisions). The reason would be due to the capacity expansion by opening a new pipeline project that connects Greece and Bulgaria. In addition, the natural gas consumption of Croatia, Sweden, and Denmark are also higher in Case 3 than those prescribed in Case 2. In contrast, the consumption of Lithuania is lower in Case 3 than the one in Case 2. Our results show that the remainder consumers are not affected significantly under the given set of infrastructure decisions.

Figure 3.4 shows the effect of infrastructure decisions on purchase price of consumer countries. In particular, the percentage change in purchase price of each consumer country is shown in a map in Figure 3.4. This percentage is calculated by first taking difference between monthly natural gas purchase price with and without the infrastructure decisions by monthly purchase price without infrastructure decisions, then averaging all these percentage differences across five year for each consumer country. Note that color coding is given with respect to the percentage change in purchase price, as detailed in Figure 3.4, and countries highlighted in grey are out of scope for this paper. Results show that natural gas purchase price of Bulgaria, Denmark, Croatia, and Sweden are significantly lower in Case 3 when compared to those in Case 2. On the other hand, that of Lithuania is higher in Case 3 compared to that in Case 2. Overall,

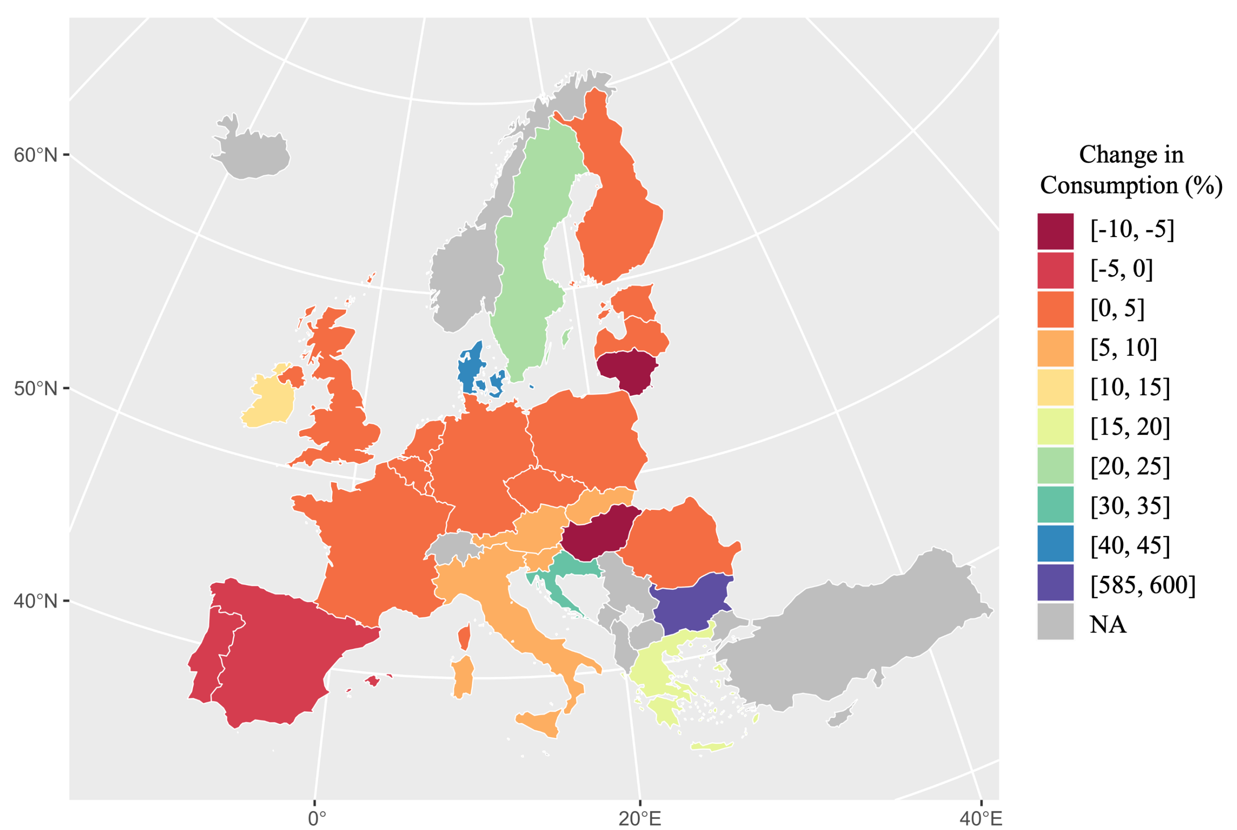


Figure 3.3 The effect of investments on consumption of 26 EU countries

Figures Figure 3.3 and Figure 3.4 reveal that new infrastructural investments yield higher consumer welfare for Bulgaria, Denmark, Croatia, and Sweden, while they yield a lower consumer welfare for Lithuania.

Results show that optimal investment decisions suggest in opening all four LNG regasification terminals and 11 out of 15 pipelines. Specifically, in B, Table B.2 and Table B.1 display timing of new investment decisions. Note that timing of opening each project varies in five years due to the annual budget of 2,000 million €.

3.5.2.3 Excluding Russian Gas Supply from the Market

In this section, we analyze the European natural gas network model using two case studies in order to examine the effect of excluding Russian gas supply from the market with and without infrastructure decisions. To this end, we consider two cases, namely Case 4 and Case 5, both excluding Russian gas supply. Particularly, Case 5 considers infrastructure decisions along with trade and storage decisions of all supply chain components, while Case 4 only considers trade and storage decisions without taking into consideration infrastructure decisions. Note that both cases consider a time horizon that spans five years between 2020-2025.

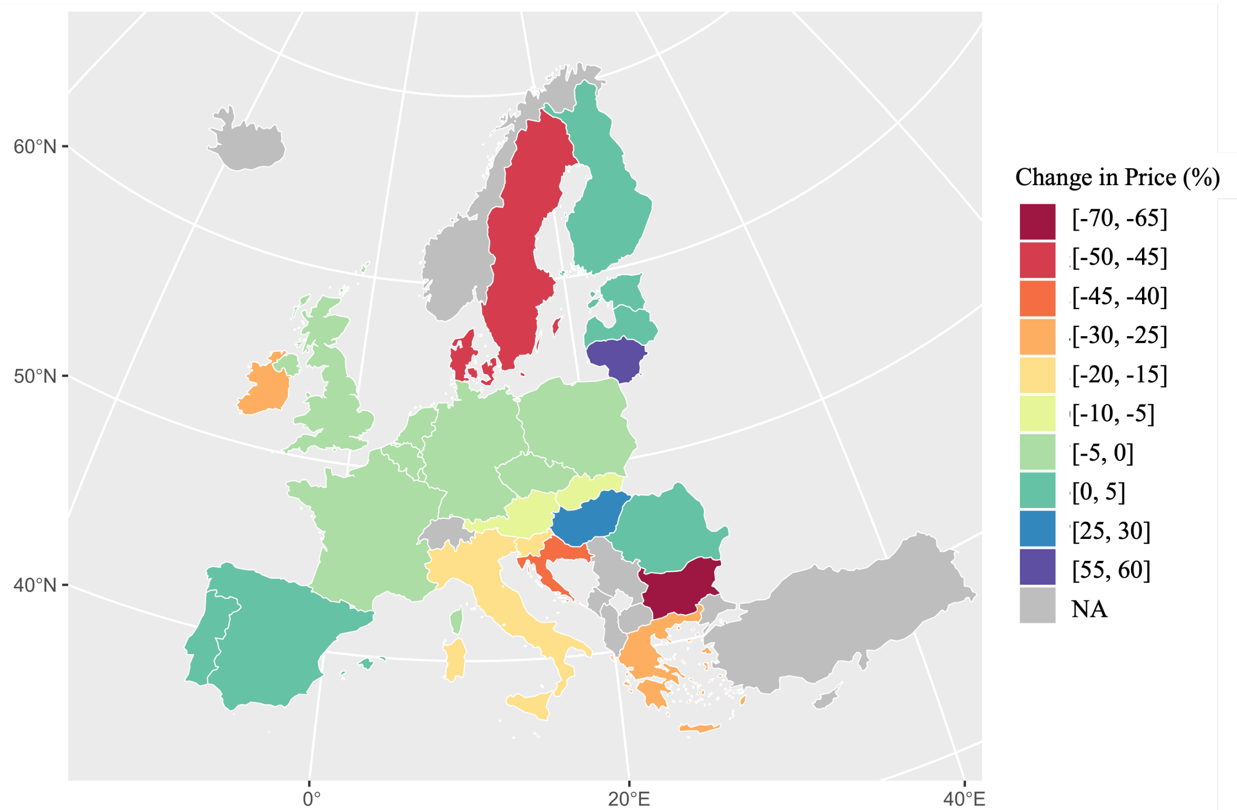


Figure 3.4 The effect of investments on purchase price of 26 EU countries

The proposed model prescribes the maximum social welfare of 121,572 million for Case 4, where only current pipeline and LNG capacities are considered without the inclusion of infrastructure decisions. In contrast, the maximum social welfare prescribed for Case 5 is 124,667 million, where the infrastructure decisions are also taken into account with 2,000 million € /year budget. In addition, we compare Case 2 with Case 4, and Case 3 with Case 5 in order to examine the effect of excluding Russian gas supply from the market. It reveals that the maximum social welfare of 26 EU countries during five years is about 10% lower in Cases 4 and 5 (without Russian gas supply) compared to those in Cases 2 and 3 (with Russian gas supply), respectively. We also calculate the value of country-level HHI for extraction stage based on the optimal decisions of Cases 4-5 prescribed by the proposed model. In particular, HHI is 2,036 for Case 4, while that for Case 5 is 2,017. This shows that in Cases 4 and 5, where the Russian natural gas supply is excluded from the market, Norway and Qatar fill this supply gap; therefore, the optimal solutions suggest that Norway and Qatar have the largest market share.

Comparing the optimal solutions of Cases 4 and 5 shows that the total available natural gas is increased from 2,224,639 mcm to 2,333,209 mcm if potential infrastructure investments are considered. In particular, the change of total natural gas consumption of each consumer country without and with infrastructure decisions are shown in a map in Figure 3.5 (a) and Figure 3.5 (b), respectively. This change is calculated by first taking difference between monthly natural gas consumption with and without Russian gas supply, then averaging all these differences across five year for each country. Note that color coding is given with respect to the difference in consumption, as detailed in Figure 3.5, and countries highlighted in grey are out of scope for this paper. Results show that natural gas consumption of Germany and Romania in Case 4 are 532 mcm and 385 mcm lower than those in Case 2. Moreover, natural gas consumption of these countries in Case 5 are 651 mcm and 335 mcm lower than those in Case 3 as shown in Figure 3.5. Comparing Figure 3.5 (a) with Figure 3.5 (b) reveals that the consumption difference for Estonia, Luxembourg, Ireland, and Croatia are lower in Figure 3.5 (b) which implies that the infrastructure decisions reduces this sensitivity, therefore the supply risk of consumer countries.

Figure 3.6 shows the effect of exclusion of Russian gas supply on purchase price of consumer countries. In particular, the logarithmic difference in purchase price of each consumer country without and with infrastructure decisions are shown in a map in Figure 3.6 (a) and Figure 3.6 (b), respectively. This difference is calculated by first taking difference between monthly natural gas purchase price with and without Russian gas supply, then averaging all these differences across five year for each consumer country. Note that color coding is given with respect to the logarithmic difference in purchase price, as detailed in Figure 3.6. Moreover, countries highlighted in grey excluding Bulgaria are out of scope for this paper. Note that, Bulgaria is also highlighted in grey since the average difference on purchase price is calculated as zero. Results show that natural gas purchase price of Finland, Latvia, and Estonia are significantly lower in Case 4 compared to that in Case 2 as shown in Figure 3.6 (a). On the other hand, those for Greece, and Lithuania are calculated nearly zero. Comparing Figure 3.6 (a) with Figure 3.6 (b) yields that the purchase price difference for Estonia are lower in Figure 3.6 (b) implying that the infrastructure decisions reduces this sensitivity.

3.6 Conclusion

This paper examines the European natural gas supply chain as a network model that includes producers and consumers, where consumer countries may include a mid-streamer, a storage, and end-users from three sectors, namely, household, industry, and power sectors. For this problem, we propose a mixed-integer non-linear programming formulation that prescribes optimal export decisions of producers and import decisions of mid-streamers via long-term contracts or spot markets, and storage decisions of mid-streamers as well as infrastructure decisions of the regulator in order to maximize the social welfare. To the best of our knowledge, this is the first study that jointly considers all these decisions along with endogenous demand at monthly-level and that examines the effect of supply risk of producers.

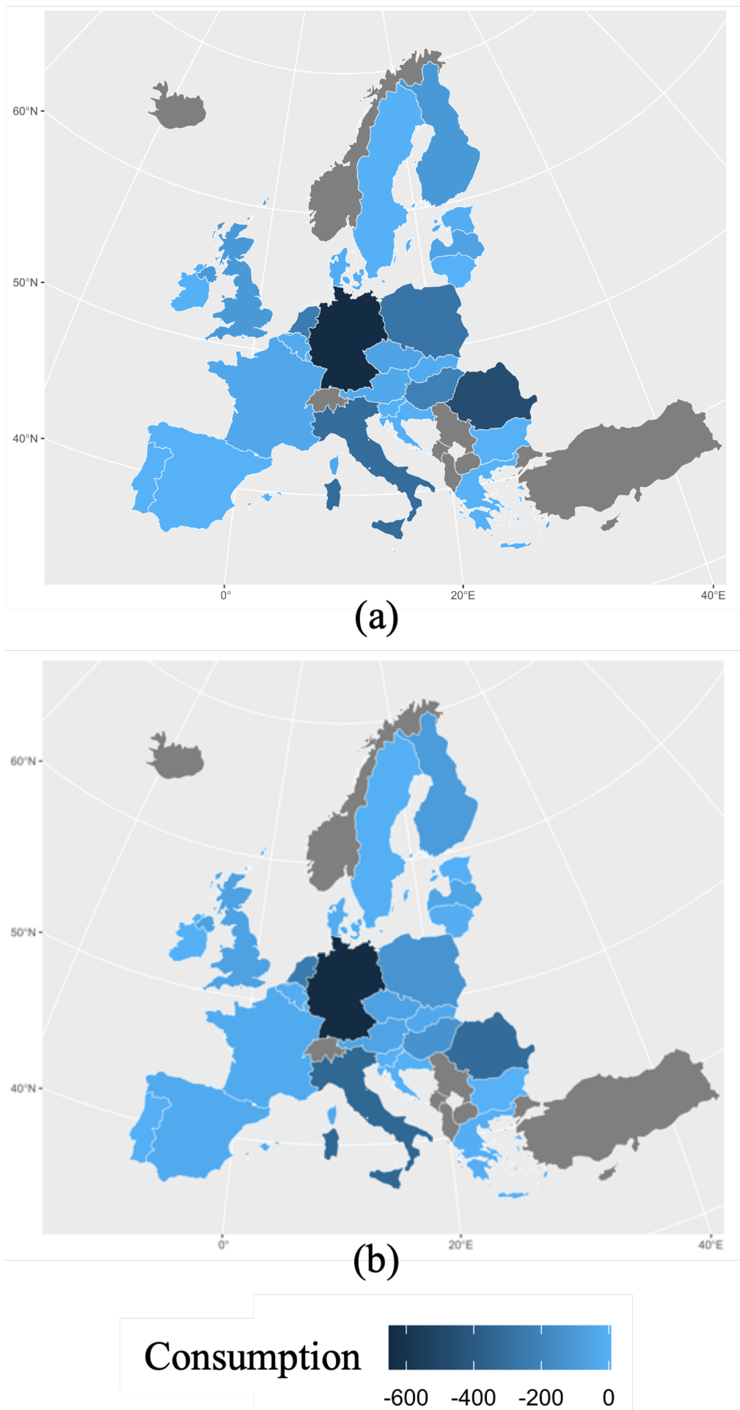


Figure 3.5 The map of the effect of Russian gas interruption on consumption without (a) and with (b) new infrastructure investments for 26 EU countries

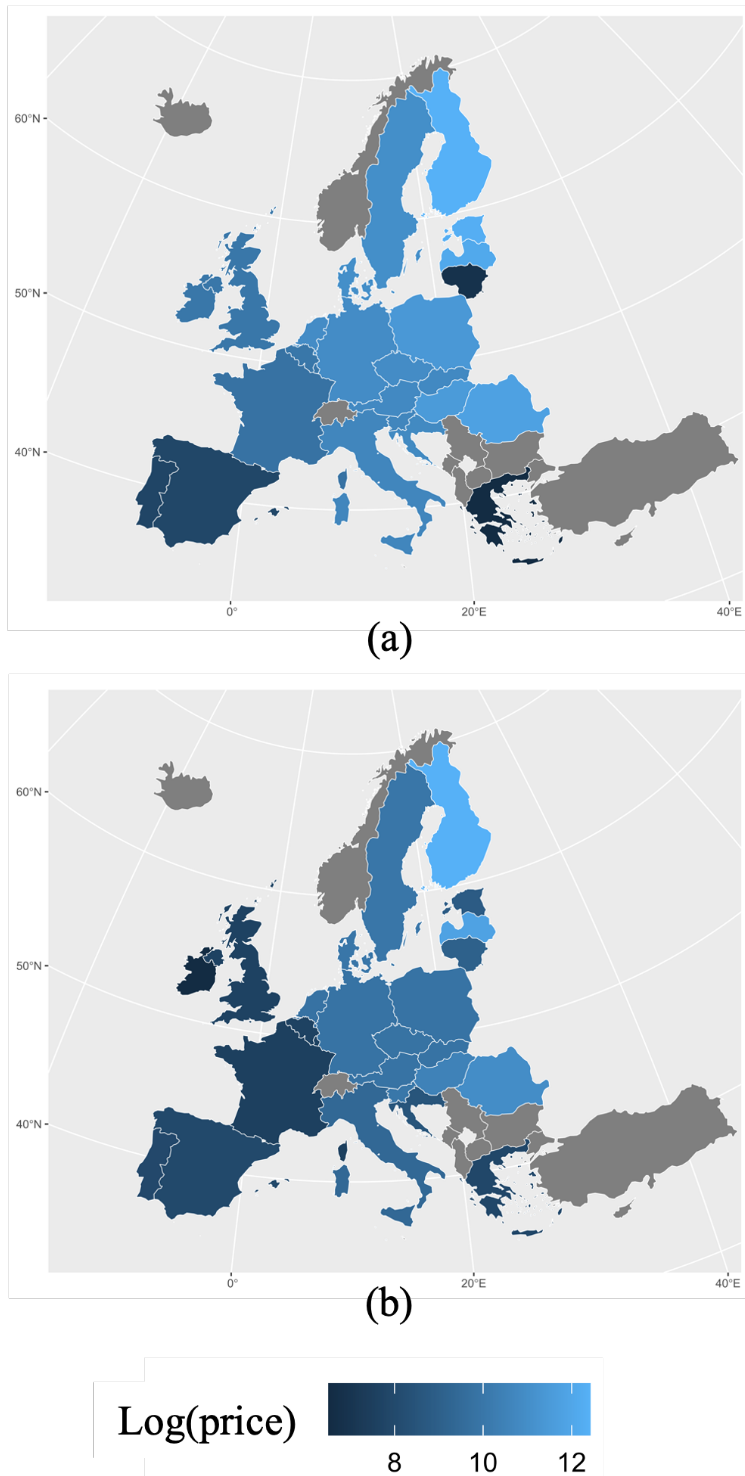


Figure 3.6 The map of the effect of Russian gas interruption on price without (a) and with (b) new infrastructure investments for 26 EU countries

In this paper, we conduct several case studies to analyze the problem under different conditions. First, we compare the optimal demand volume of each end-user for each sector with actual consumption volumes of these European countries during 2019. Secondly, we implement the proposed model with and without infrastructure decisions. Comparing these two cases reveals that expanding capacities with new infrastructure investments increases the total social welfare by nearly three billion for 26 EU countries. In particular, these investments yield higher consumer welfare for Bulgaria, Denmark, Croatia, and Sweden, while they yield a lower consumer welfare for Lithuania. In addition, optimal investment decisions suggest in opening all four LNG regasification terminals and 12 out of 16 pipelines.

We also compare several cases by excluding Russian gas supply from the EU natural gas market with and without infrastructure decisions. Our results suggest that, if Russian gas supply is excluded from the market, than the social welfare and cumulative natural gas consumption of 26 EU countries decrease by 10% and 15%, respectively. Specifically, the consumption difference for Germany and Romania are highest among other countries. Moreover, infrastructure investments can reduce those differences in consumption for Estonia, Luxembourg, Ireland, and Croatia. Moreover, results reveal that if there is no infrastructure investments, the purchase price of natural gas for Finland, Latvia and Estonia are mostly affected among other EU countries by exclusion of Russian gas supply. In addition, if there are infrastructure investments, then results reveal that the difference in purchase price for Estonia decreases by opening a new regasification terminal. Therefore, making investments on LNG terminals and pipelines would reduce supply risk of consumer countries.

In this study, we consider a given set of possible infrastructure investments to determine the optimal ones to be opened in the model. However, there may exist other possible locations where potential pipelines and LNG regasification terminals can be constructed. Therefore, as future work, one may consider to alter the proposed model such that it prescribes infrastructure decisions considering all these alternatives.

CHAPTER 4

HOW TO MEET LITHIUM REQUIREMENT FOR THE NEXT 20 YEARS? AN OPTIMIZATION MODEL OF GLOBAL LITHIUM CLOSED-LOOP SUPPLY CHAIN

Targeted for *Resources, Conservation* & *Recycling.*

Merve Olmez Turan[1,2], Roderick Eggert[4], Tulay Flamand[3]

4.1 Abstract

With the increasing trend of clean energy transition, the global lithium requirement is expected to increase continuously. However, current raw material production capacity is insufficient to meet the expected demand. We address the global closed-loop lithium supply chain with several actors, such as brine, hard rock mines and recycling facilities for nine regions which cover the global. We propose a mixed-integer linear programming model which prescribes decisions regarding the amount of lithium-carbonate traded among these supply chain actors as well as investment decisions for new brine and hard rock deposits in order to minimize the total cost, including direct cost and cost of carbon from 2019 to 2040. We also examine the sensitivity of the model to the changes in the social cost of carbon (SCC) and locations of cathode manufacturers. Our results suggest that the cumulative cost of supplying lithium carbonate for this period is $ $157,083$ million. Moreover, recycling has a significant role in the market after 2036, while the market shares of brine and hard rock operations are similar without considering SCC. However, SCC prioritizes hard rock or brine operations with or without additional cathode manufacturers. In addition, it decreases the cumulative global warming potential (GWP) by 53%, or 31% depending on the cathode manufacturers' location.

4.2 Introduction and Literature Review

Lithium has been used in several sectors such as ceramic, steel, and glass, since it is the lightest and most conductive metal in the Earth's crust. Moreover, lithium plays a crucial role in the global transition to a low-carbon economy since it cannot be substituted in Lithium-ion batteries (LIB) which is one of the primary technology in this transition [63]. To this end, lithium demand has significantly increased during the past ten years. Specifically, in 2013, global lithium demand was nearly 160,000 metric tonnage (t) lithium carbonate equivalent (LCE), whil six years later, it nearly doubled, and more than half of them

[1]Graduate student, Colorado School of Mines
[2]Primary researcher and author
[4]Professor, Colorado School of Mines
[3]Assistant Professor, Colorado School of Mines

were consumed for LIB, specifically for electrical vehicles (EVs) [64]. Since the global transition from internal combustion engine vehicles to EVs has just started, the lithium consumption growth is expected to increase at a higher rate than it has over the past six years [65]. According to Benchmark Mineral Intelligence [65], capacities of the potential primary and secondary raw material producers (RMP) are expected to be insufficient to meet global lithium demand after 2025. Moreover, Sonoc et al. [66] suggest that in the near future, the world would be facing a severe scarcity of lithium without recycling of LIB. Moreover, the recent studies support that the recycling end-of-life (EOL) LIB may be a solution for lithium scarcity [67, 68].

Deployment of EVs is prioritized by governments in order to reduce global CO_2 emissions in transportation sectors. Specifically, transportation caused 37% of the global CO_2 emissions from end-user sectors in 2018 [69]. EVs can reduce CO_2 emissions depending on electricity supply sources [70]. For example, if their sourced electricity does not cause any CO_2 emissions, their well to wheel CO_2 emissions, which the summation of CO_2 emissions caused by all process in order to travel a mile, would be zero. However, the global warming potential (GWP) of EV manufacturing is higher than that of internal combustion engine vehicles, with an additional 13 grams of carbon dioxide equivalent per kilometer for LIB manufacturing. These emissions includes emissions caused by battery material extraction, energy and material requirement for LIB manufacturing [71]. Zero-carbon electricity production can reduce the emissions caused by energy requirements of manufacturing and transportation. However, the GWP of mineral extraction should also be considered in order to achieve the net-zero carbon target.

In this study, we aim to answer the following research questions: (i) What are the optimal decisions of the selection of primary and secondary RMPs, their annual capacities and production levels, and the amount of lithium transported between all supply chain components that minimize the total cost (i.e., capital costs, operational and transportation costs and the cost of carbon) while meeting the future lithium requirement? (ii) How sensitive these decisions and the GWP of the supply chain are to the changes in the social cost of carbon (SCC)? (iii) How do additional cathode manufacturers in the United States (US) and Europe affect the extraction and recycling decisions of the lithium supply chain?

To answer these questions, we propose a mixed-integer linear programming model that prescribes decisions for the extraction, chemical transformation, consumption and recycling stages of the lithium supply chain for nine regions that span the entire world in order to minimize the total cost. Note that since our study aims to investigate extraction and recycling operations, we do not consider intermediate goods manufacturers stage which includes battery and final goods manufacturing. In the literature, several papers investigate the LIB supply chain, while there are fewer studies specifically focusing on the lithium market [72–74]. In this regard, our work is most related to Rosendahl and Rubiano [75] who investigate the closed-loop lithium supply chain by employing a system dynamics model, considering the recycling effect on lithium scarcity and market intervention. By contrast, we examine the closed-loop lithium supply chain under considerations of the SCC and additional cathode manufacturers on extraction and recycling stages. We also investigate deposit-by-deposit primary RMPs in order to gain insights regarding the market concentration and country-level supply risk.

In the literature, there are limited studies that examine the lithium supply chain. Specifically, some studies propose system dynamics and material flow analysis in order to investigate the recovery of lithium from EOL LIB effects under different energy scenarios [76–78]. Sverdrup et al. [76] the first system dynamics model, namely LITHIUM, in order to estimate the lithium market price with endogenous feedback on mining, demand, and recycling. The study suggests that improvement in recycling and limitation on irreversible lithium losses in 2015–2025 potentially prevents lithium scarcity before 2100 [76]. Moreover, Li et al. [79] propose the first closed-loop supply chain network model for the EV LIB industry that considers LIB remanufacturing infrastructure, providing that a 30.93% increase in profit can be achieved by integrated remanufacturing infrastructure into standalone EV LIB manufacturing supply chains. To the best of our knowledge, our contribution to the literature is the following: (i) We propose the first mixed-integer programming formulation that jointly considers production, import, export, and stock decisions of lithium supply chain components that minimize the total capital, operational, and transportation costs, (ii) we also investigate the SCC effect on these decisions for the global lithium market, (iii) it is also the first study that examines the effect of market concentration of cathode manufactures on these decisions as well as GWP of the lithium supply chain.

The remainder of this paper is organized as follows: Section 4.3 introduces a formal problem statement with our model assumptions, notation, and the proposed optimization model. Section 4.4 provides our computational study with the details of the data acquisition and computational results. This section also conducts a sensitivity analysis for SCC and cathode manufacturers in US and Europe. Finally, Section 4.5 concludes the paper with a discussion of future work.

4.3 Methodology

In this section, we provide a formal problem statement for the closed-loop lithium supply chain, which motivates the development of a mixed-integer linear programming model, also proposed in this section.

4.3.1 Problem Statement

In this paper, we examine the closed-loop supply chain as a network model in which the nodes represent several supply chain actors. These actors include primary and secondary RMPs (i.e., brine, hard rock mine, and recycling facility), mineral conversion plants, cathode manufacturers, battery and final product manufacturers, and end-users, while directed arcs between nodes represent material flow from one another.

A sample network representation of the closed-loop lithium supply chain is shown in Figure 4.1. Lithium is primarily extracted either from brine or hard rock mines. Note that although Figure 4.1 includes only one node for hard rock and brine sources, we consider several of them, one for each deposit. Brine-sourced lithium is converted to lithium carbonate in the same region where the brine deposit is located. Then, lithium carbonate may be sent directly to end-users for non-LIB sector or cathode manufacturers in order to produce cathode powder. Note that nine of these non-LIB sub-sectors (e.g., ceramic, glass, grease) are considered jointly in one node since lithium cannot be recoverable from the non-LIB products. Moreover, end-users of non-LIB sectors are also considered as manufacturers of these nine sectors. For example, ceramic manufacturers in China are considered as a part of the non-LIB end-user for China. However, these manufacturers may send their products to the ceramic customers in the US, whom we do not consider. As opposed to brine-sourced lithium, hard rock-sourced lithium is mainly in spodumene form. Concentrated spodumene is sent to mineral conversion plants which are mainly located in China for converting it to lithium carbonate. All of the mineral conversion plants' products may either be sent to end-users for non-LIB or cathode manufacturers. We assume that lithium is not recoverable from non-LIB products. To this end, when the lithium products for non-LIB sectors reach their EOL, they are sent to waste fills which are not considered in this study. On the other hand, brine and hard rock-sourced lithium is used to produce cathode powders in cathode manufactures.

All cathode powders which include lithium are sent to the battery and final product manufacturers in order to produce LIBs and final products (e.g., EVs, computers, e-bikes). Although there are multiple battery and final product manufacturers all over the world, we consolidate all of them into a single node for connecting cathode manufactures and LIB end-users (i.e., EVs, energy storage systems, and other electronics) since battery and final product manufacturers do not have direct connection with primary or secondary RMPs. To this end, they do not have any economic and environmental relationship. With this assumption, we exclude the cost of transportation for the routes which include connections from cathode manufacturers to battery and final product manufacturers as well as those from battery and final product manufacturers to LIB end-users. We also assume that the capacity for producing the battery and final product is sufficient in order to meet future requirements for products that are made of lithium. In this study, the supply chain analysis specifically aim to investigate raw material production shares

As shown in Figure 4.1, in LIB sector, we consider three LIB sub-sectors as highlighted in blue nodes, namely, EVs, energy storage systems (ESSs), and other electronics since the LIB for each sector has a different life time, which plays significant role for EOL battery flows over years. However, instead of product manufactures, LIB end-user nodes represent the final end-users such as EV driver and laptop users since lithium can be recoverable from LIBs after reaching their EOL. The final products where LIBs are embedded coming from battery and final good manufacturers may be sent to end-users of these three LIB sectors in all regions.

When the LIB products complete their life time, they are collected in EOL battery stocks which is located in the same region with the end-user of the LIB. When recycling of LIB is economically and environmentally feasible, then the required amount of EOL LIB may also be sent to recycling facilities to recover lithium. Note that the collection, storage, and transportation cost of EOL battery and capital expenditures (CAPEX) are assumed to be zero, since the lithium is considered as a by-product of LIB recycling operations because the main purpose of LIB recycling facilities is to recover Cobalt and Nickel from EOL batteries. Finally, all recovered lithium is sent to cathode manufacturers. In this study, we only consider the hydro metallurgical process as LIB recycling, while we do not consider direct-recycling and pyro metallurgical processes of lithium recovery.

In our study, we also consider the geography of these nodes in the lithium supply chain. These nodes are located in some or all nine regions, namely, Pacific (PAC), East Asia (EAS), Europe and Central Asia (ECA), Latin America and the Caribbean (LAC), Middle East and North Africa (MEN), North America (NAM), South Asia (SAS), Sub-Saharan Africa (SSA), and China, which is discussed in more detail in Section 4.4.1. Since China has the dominant position in the global lithium supply chain with respect to production capacities of mines and mineral conversion plants and cathode manufacturers, we consider it as a standalone region, separate from EAS. Specifically, 22% of global raw material capacities, 99% of global mineral conversion plant capacities, and 72% of global cathode manufacturer capacities belong to China [64].

Our study examines this supply chain to prescribe optimal decisions in order to minimize the total supply chain costs which also include the cost of carbon emissions. We later conduct a sensitivity analysis to investigate the effect of this cost on the optimal decisions. Note that our study examines carbon emissions in the context of three scopes provided by the US Environmental Protection Agency (EPA). Specifically, Scope 1 includes emissions that are caused by an organization's activity directly. By contrast, Scope 2 covers indirect emissions associated with energy purchased in order to complete the organization's activities. Finally, Scope 3 consists of those that occur in each supply chain activity, including those in Scope 1 and Scope 2. Figure 4.2 lists what we consider from each scope in the EPA's framework in our study. Specifically, in Scope 1 and Scope 2 emissions, we consider those related to mineral conversion plants, primary RMPs, and secondary RMPs. Moreover, in Scope 3, we consider emissions coming from activities for supplying the materials required for RMPs and those caused by the transportation of lithium

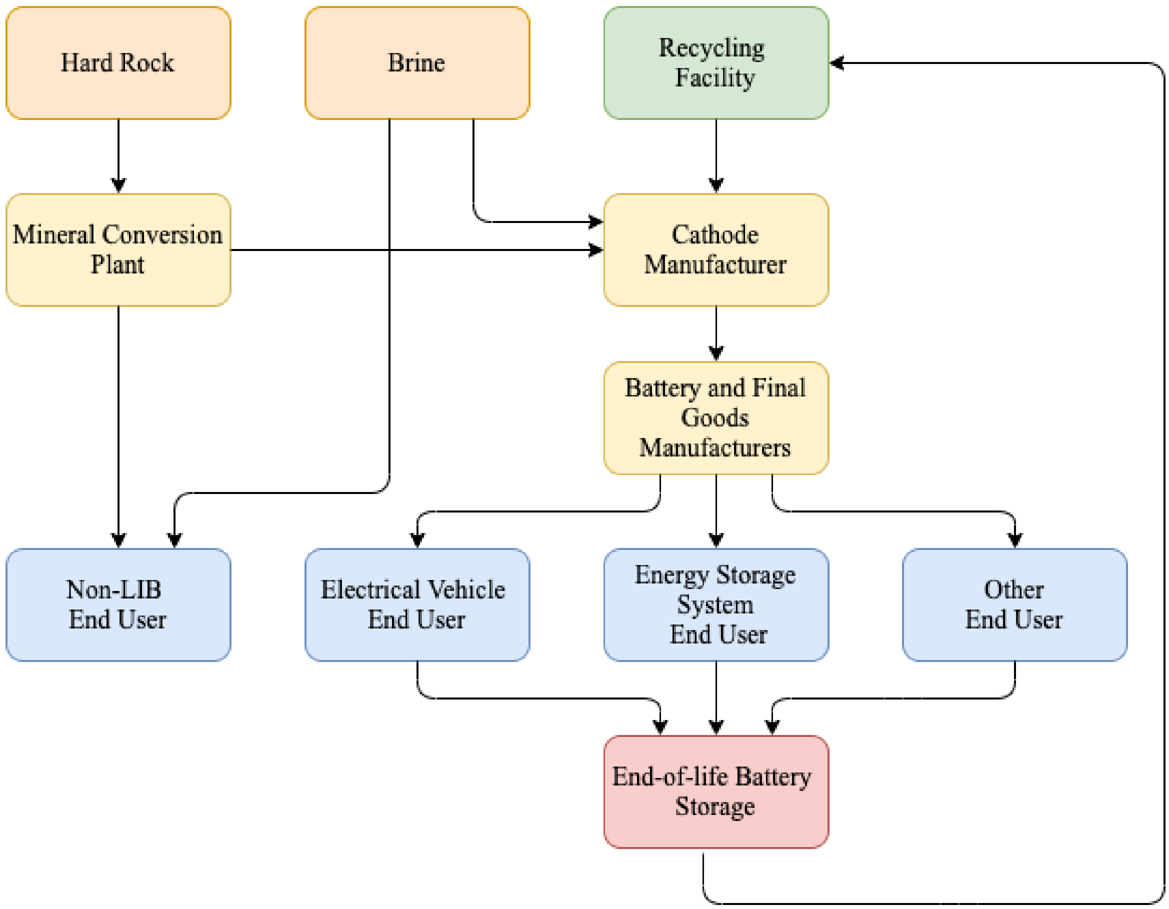


Figure 4.1 The representation of the closed-loop lithium supply chain

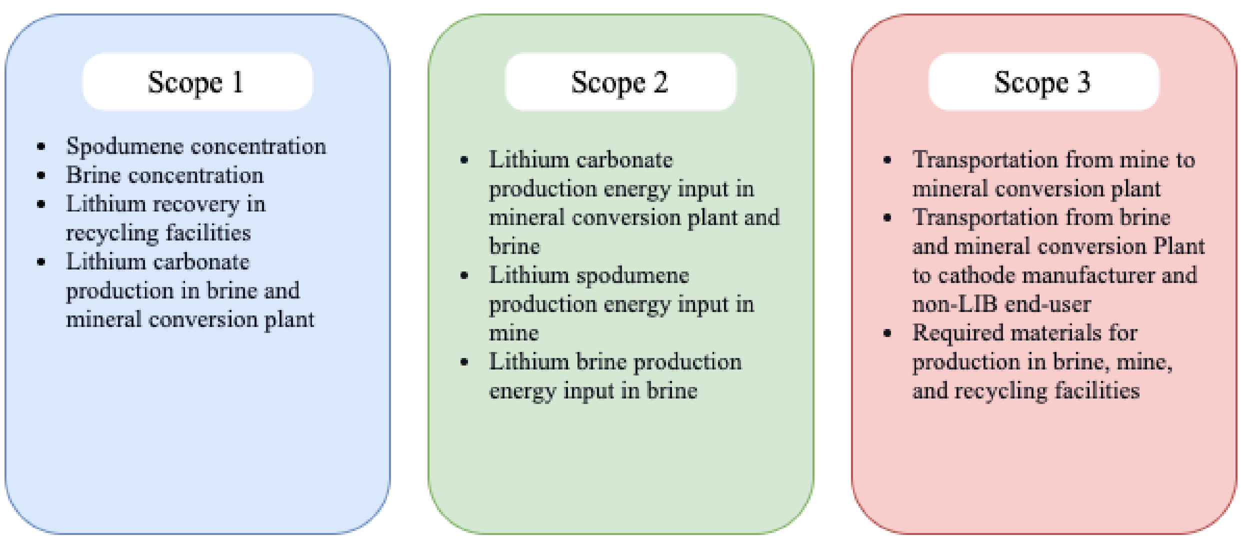


Figure 4.2 Our scope of the emission in relation to the EPA's framework

between supply chain actors.

4.3.2 Mathematical Formulation

In this section, we examine the lithium closed-loop supply chain by proposing a mixed-integer linear programming model in order to minimize the discounted cost which consists of operational expenditure (OPEX), CAPEX, transportation costs, and cost of carbon for the time horizon from 2021 to 2040 with an annual resolution. We develop this model to prescribe optimal infrastructure decisions of opening new brine and hard rock mines, production amounts and capacities for existing ones, and the lithium flow between supply chain actors. The model also provides optimal recovery amounts of recycling facilities and storage decisions of EOL LIBs in the lithium closed-loop supply chain.

Our notations and the proposed mathematical formulation, Closed-loop Lithium Model **(CLLM)**, are introduced below. Note that, sets and parameters are input for an optimization models, while variables are (unknown) decisions to be optimized. Consider the sets and parameters as follows:

Sets

$\mathcal{G}$	Set of existing and possible primary RMPs (i.e., brine and hard rock mines)	
$\mathcal{G}^{\mathrm{e}} \subseteq \mathcal{G}$	Set of existing primary RMPs	
$\mathcal{G}^{\mathrm{p}} \subseteq \mathcal{G}$	Set of possible primary RMPs	
$\mathcal{R}$	Set of secondary RMPs (i.e., recycling facilities)	
$\mathcal{L}$	Set of intermediate product manufacturers (mineral conversion plants, cathode manufacturers, battery and final good manufacturers)	
$\mathcal{E}$	Set of end-users	
$\mathcal{E}^{\mathrm{n}} \subseteq \mathcal{E}$	Set of non-LIB end-users	
$\mathcal{E}^{\mathrm{b}} \subseteq \mathcal{E}$	Set of LIB end-users (EVs, ESSs, and other electronics)	
$\mathcal{S}$	Set of EOL LIB storages	
$\mathcal{F}$	Set of secondary RMP efficiencies	
$\mathcal{A}$	Set of all arcs (i.e., connections)	
$\mathcal{T}$	Set of years	

Cost Parameters

c_g^{OPP}	Initial unit OPEX of possible or existing primary RMP g	[$/ t LCE]
c^{OPS}	Initial unit OPEX of a secondary RMP	[$/ t LCE]
c_g^{CA}	Unit CAPEX of primary RMP g capacity	[$/ t LCE]
c^{TR}	Unit transporting cost of material	[$ / mile-t LCE]
c_t^{SCC}	Unit social cost of carbon at year t	[$/ t CO_2]
c_f^{EFC}	Additional cost of a secondary RMP based on efficiency f	[$/ t LCE]

Capacity Parameters

δ_g^{R}	Amount of lithium reserve of primary RMP g	[t LCE]
δ_g^{P}	Initial lithium production capacity of possible or existing primary RMP g	[t LCE]
δ^{M}	Maximum lithium production capacity that a primary RMP can reach	[t LCE]

Demand Parameters		
d_{et}	Amount of lithium requirement of end-user e at year t	[t LCE]
d_{et}^{EOL}	Amount of lithium embedded in EOL LIB of end-user e at year t	[t LCE]

Time Dependent Parameters		
α_t^{T}	Percentage of operational expenditure reduction for a primary or secondary RMP at year t	[%]
α_t^{D}	Discount rate at year t	[%]

Carbon Emission Parameters		
n_g^{PP}	Unit carbon emission caused by production of primary RMP g	[t CO_2/t LCE]
n^{SP}	Unit carbon emission caused by recovery of a secondary RMP	[t CO_2/t LCE]
n^{TR}	Unit carbon emission caused by material transportation	[t CO_2/mile-t LCE]

Other Parameters		
a_{ij}^{DST}	Distance from region i to region j	[mile]
a_f^{EFN}	Efficiency f of a secondary RMP	[%]
a	Percentage limit of annual capacity expansion and shrinkage for a primary RMPs	[%]
M_t	Cumulative available EOL LIB from initial time period until year t	[unitless]

The decision variables of the CLLM are introduced below:

Binary Variables		
z_{gt}^{PR}	1 if possible primary RMP g is operating at year t; otherwise 0	$\{0, 1\}$
z_f^{SR}	1 if efficiency f is chosen for secondary RMP; otherwise 0	$\{0, 1\}$
Continuous Variables		
x_{ijt}	Amount of transported lithium from node i to node j at year t	[t LCE]
q_{st}	Amount of lithium which is embedded in EOL LIB at storage s at year t	[t LCE]
y_{gt}	Annual lithium production capacity for possible and existing primary RMP g at year t	[t LCE]
w_{rtf}	The amount of lithium received by the secondary RMP r at year t with efficiency f, if f is selected, 0 otherwise	[t LCE]

Model **CLLM** is then formulated as the following mixed-integer programming model:

CLLM: Minimize

$$\sum_{(g,j)\in\mathcal{A}:g\in\mathcal{G}} \sum_{t\in\mathcal{T}} \alpha_t^{\mathrm{D}} \cdot \alpha_t^{\mathrm{T}} \cdot c_g^{\mathrm{OPP}} \cdot x_{gjt} \tag{4.1a}$$

$$+ \sum_{(r,j)\in\mathcal{A}:r\in\mathcal{R}} \sum_{t\in\mathcal{T}} \alpha_t^{\mathrm{D}} \cdot \alpha_t^{\mathrm{T}} \cdot c_r^{\mathrm{OPS}} \cdot x_{rjt} \tag{4.1b}$$

$$+\sum_{r\in\mathcal{R}}\sum_{f\in\mathcal{F}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot\alpha_t^{\mathrm{T}}\cdot c_f^{\mathrm{EFC}}\cdot a_f^{\mathrm{EFC}}\cdot w_{rtf} \tag{4.1c}$$

$$+\sum_{g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot c_g^{\mathrm{CA}}\cdot y_{gt} \tag{4.1d}$$

$$+\sum_{(i,j)\in\mathcal{A}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot a_{i,j}^{\mathrm{DST}}\cdot c^{\mathrm{TR}}\cdot x_{ijt} \tag{4.1e}$$

$$+\sum_{(g,j)\in\mathcal{A}:g\in\mathcal{G}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot c_t^{\mathrm{SCC}}\cdot n_g^{\mathrm{PP}}\cdot x_{gjt} \tag{4.1f}$$

$$+\sum_{(r,j)\in\mathcal{A}:r\in\mathcal{R}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot c_t^{\mathrm{SCC}}\cdot n^{\mathrm{SP}}\cdot x_{rjt} \tag{4.1g}$$

$$+\sum_{(i,j)\in\mathcal{A}}\sum_{t\in\mathcal{T}}\alpha_t^{\mathrm{D}}\cdot a_{i,j}^{\mathrm{DST}}\cdot c_t^{\mathrm{SCC}}\cdot n^{\mathrm{TR}}\cdot x_{ijt} \tag{4.1h}$$

$$\sum_{(g,j)\in\mathcal{A}}\sum_{t\in\mathcal{T}}x_{gjt}\leq\delta_g^{\mathrm{R}}, \qquad \forall g\in\mathcal{G} \tag{4.2a}$$

$$\sum_{(g,j)\in\mathcal{A}}x_{gjt}\leq y_{gt}, \qquad \forall g\in\mathcal{G},t\in\mathcal{T} \tag{4.2b}$$

$$y_{gt}\leq\delta^{\mathrm{M}}, \qquad \forall g\in\mathcal{G}^{\mathrm{e}},t\in\mathcal{T} \tag{4.2c}$$

$$y_{g0}=\delta_g^{\mathrm{P}}, \qquad \forall g\in\mathcal{G}^{\mathrm{e}} \tag{4.2d}$$

$$\sum_{g\in\mathcal{G}^{\mathrm{p}}}y_{g0}=0, \tag{4.2e}$$

$$y_{gt}\geq(1-a^{\mathrm{LMT}})\cdot y_{g(t-1)}, \qquad \forall g\in\mathcal{G}^{\mathrm{e}},t\in\mathcal{T} \tag{4.2f}$$

$$y_{gt}\leq(1+a^{\mathrm{LMT}})\cdot y_{g(t-1)}, \qquad \forall g\in\mathcal{G}^{\mathrm{e}},t\in\mathcal{T} \tag{4.2g}$$

$$y_{gt}\leq\delta^{\mathrm{M}}\cdot z_{gt}^{\mathrm{PR}}, \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2h}$$

$$y_{gt}\geq\delta_g^{\mathrm{P}}\cdot(z_{gt}^{\mathrm{PR}}-z_{g(t-1)}^{\mathrm{PR}}), \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2i}$$

$$y_{gt}\leq\delta_g^{\mathrm{P}}+\delta^{\mathrm{M}}\cdot(1-z_{gt}^{\mathrm{PR}}+z_{g(t-1)}^{\mathrm{PR}}), \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2j}$$

$$y_{gt}\geq(1-a^{\mathrm{LMT}})\cdot y_{g(t-1)}-\delta^{\mathrm{M}}\cdot(z_{gt}^{\mathrm{PR}}-z_{g(t-1)}^{\mathrm{PR}}), \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2k}$$

$$y_{gt}\leq(1+a^{\mathrm{LMT}})\cdot y_{g(t-1)}+\delta^{\mathrm{M}}\cdot(z_{gt}^{\mathrm{PR}}-z_{g(t-1)}^{\mathrm{PR}}), \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2l}$$

$$z_{g(t-1)}^{\mathrm{PR}}\leq z_{gt}^{\mathrm{PR}}, \qquad \forall g\in\mathcal{G}^{\mathrm{p}},t\in\mathcal{T} \tag{4.2m}$$

$$\sum_{g\in\mathcal{G}^{\mathrm{p}}}z_{g0}^{\mathrm{PR}}=0, \tag{4.2n}$$

$$\sum_{(i,e)\in\mathcal{A}}x_{iet}\geq d_{et}, \qquad \forall e\in\mathcal{E},t\in\mathcal{T} \tag{4.2o}$$

$$\sum_{(i,l)\in\mathcal{A}}x_{ilt}=\sum_{(l,k)\in\mathcal{A}}x_{lkt}, \qquad \forall l\in\mathcal{L},t\in\mathcal{T} \tag{4.2p}$$

$$q_{st}=q_{s(t-1)}+\sum_{(i,s)\in\mathcal{A}}x_{ist}-\sum_{(s,k)\in\mathcal{A}}x_{skt}, \qquad \forall s\in\mathcal{S},t\in\mathcal{T} \tag{4.2q}$$

$$q_{s0}=0, \qquad \forall s\in\mathcal{S} \tag{4.2r}$$

$$\sum_{(e,j)\in\mathcal{A}}x_{ejt}=d_{et}^{\mathrm{EOL}}, \qquad \forall e\in\mathcal{E}^{\mathrm{b}},t\in\mathcal{T} \tag{4.2s}$$

$$\sum_{(i,r)\in\mathcal{A}}x_{irt}\geq w_{rtf}, \qquad \forall r\in\mathcal{R},f\in\mathcal{F},t\in\mathcal{T} \tag{4.2t}$$

$$w_{rtf} \leq M_t \cdot z_f^{\mathrm{SR}}, \qquad \forall r \in \mathcal{R}, f \in \mathcal{F}, t \in \mathcal{T} \tag{4.2u}$$

$$w_{rtf} \geq \sum_{(i,r)\in\mathcal{A}} x_{irt} - M_t \cdot (1 - z_f^{\mathrm{SR}}), \qquad \forall r \in \mathcal{R}, f \in \mathcal{F}, t \in \mathcal{T} \tag{4.2v}$$

$$\sum_{(r,j)\in\mathcal{A}} x_{rjt} \leq \sum_{f\in\mathcal{F}} a_f^{\mathrm{EFN}} \cdot w_{rtf}, \qquad \forall r \in \mathcal{R}, t \in \mathcal{T} \tag{4.2w}$$

$$\sum_{f\in\mathcal{F}} z_f^{\mathrm{SR}} = 1, \tag{4.2x}$$

$$\mathbf{x}, \mathbf{q}, \mathbf{y}, \mathbf{w} \geq 0, \mathbf{z}^{\mathrm{PR}}, \mathbf{z}^{\mathrm{SR}} binary. \tag{4.2y}$$

The objective function (4.1a)-(4.1e) minimizes the discounted total cost. Specifically, (4.1a) represents the total discounted OPEX for possible and existing primary RMPs, while (4.1b) represents that for secondary RMPs. The term (4.1c) represents an additional discounted OPEX due to the selected efficiencies for secondary RMPs. Note that we calculate the parameter c_f^{EFC} by using a cost function adapted from Rosendahl and Rubiano [75] who propose a non-linear operational cost function under a joint consideration of technology adoption and recycling facility efficiency. The detailed discussion of this function is presented in the Appendix C with the discussion of its linearization. The term (4.1e) presents the total discounted transportation cost throughout the supply chain. In addition, the term (4.1d) calculates the discounted CAPEX for primary RMPs. Terms (4.1f)-(4.1g) represent the discounted SCC caused by the production of existing and possible primary RMPs and secondary RMPs, respectively, while the term (4.1h) calculates the total discounted social cost of carbon caused by the transportation throughout the supply chain. Constraint (4.2a) enforces that the total extraction of existing or possible primary RMP g during the entire time horizon cannot exceed its total reserve. Moreover, Constraint (4.2b) ensures that the annual production of possible or existing primary RMP g cannot exceed its annual production capacity. Note that if the short time horizon, than constraint (4.2a) would be redundant. Constraint (4.2c) ensures that annual production capacity for existing primary RMPs g cannot exceed the given upper limit, whereas Constraint (4.2d) assigns the initial annual production capacities for each existing primary RMP g at the beginning of the time horizon. Constraint (4.2e) ensures that all initial capacities are zero for possible (non-existing) primary RMPs. Constraints (4.2f)-(4.2g) guarantee that the production capacity for primary RMP g at year t is determined between $\pm$ a^{LMT}% of its capacity of the previous year. Constraints (4.2h)-(4.2n) ensure capacity restrictions for non-existing primary RMPs. Specifically, Constraint (4.2h) enforces the capacity of primary RMP g at time t to be zero if RMP g is not opened for this time period; otherwise, this capacity is limited with an upper limit. Constraints (4.2i)-(4.2j) ensure that if primary RMP g is opened at year t, its corresponding production capacity must be equal to the given initial capacity for that year. Constraints (4.2k)-(4.2l) guarantee that the production capacity for primary RMP g at year t is determined between $\pm$ a^{LMT}% of its capacity of the previous year, if it is opened before year t. Constraint (4.2m) enforces that if a possible primary RMP g is opened at year $t-1$, then it must stay opened for all years after $t-1$. Constraint (4.2n) ensures that all possible primary RMPs are initially inactive. Constraint (4.2o) ensures that the total amount of lithium transported to end-user e must meet its annual lithium requirement. Constraint (4.2p) is a flow-balance constraint for intermediate product manufacturer l that guarantees that the total amount of lithium received by the manufacturer l should be equal to the total amount of lithium sent from it. Constraint (4.2q) is a storage balance constraint for each EOL storage s at year t. Particularly, the amount of lithium stored in EOL storage s at year t equals last year's available storage amount plus the total received lithium from a battery end-user e at year t minus the amount of EOL which is considered to be sent out at year t. Constraint (4.2r) enforces that there is no initial inventory for any EOL storage s. Moreover, Constraint (4.2s) enforces that the total amount of lithium sent from each LIB end-user e to its accessible EOL storage must be equal to the amount of lithium embedded in EOL LIB of this end-user e. Constraints (4.2t)-(4.2v) calculate the amount of lithium received by the secondary RMP r based on the selected efficiency f at year t. Constraint (4.2w) represents the flow-balance constraint for each secondary RMP r at year t. Specifically, it ensures that the total lithium amount sent from the secondary RMP r must be less than the total amount of recovered lithium by this secondary RMP r calculated based on the selected efficiency f. Constraint (4.2x) enforces that only one efficiency must be selected. Finally, Constraint (4.2y) enforces non-negativity and binary restrictions on decision variables.

4.4 Computational Study

In this section, we detail our data acquisition, and conduct a computational study by implementing the proposed model **CLLM** on the obtained data. Finally, we provide and discuss our results.

4.4.1 Data Set Acquisition

In this section, we provide details regarding the data obtained and used as model parameters in the proposed model **CLLM**. The proposed model considers the global lithium market as a network, including the set of primary and secondary RMPs, mineral conversion plant, cathode manufacturer, and end-user nodes. All of these nodes are located in some or all of nine regions: PAC, EAS, ECA, LAC, MEN, NAM, SAS, SSA, and China. Specifically, we consider 22 primary raw material producers; eight brine operations in LAC, NAM, and China, and 14 hard rock mines in PAC, ECA, SSA, and China. Moreover, the set of possible RMPs includes 43 brine deposits in LAC, NAM, and China and 60 hard rock deposits in nine regions except for EAS and MEN. The data set covers the countries where the brine and hard rock deposits are located. In addition, we consider the following model attributes: (i) a single node for mineral conversion plants since 99% of global mineral conversion capacity belong to China; (ii) a single node for battery and final product manufacturers since they do not significantly affect the investment decisions; (iii) two sets of cathode manufacturers which consist of two nodes in EAS and China and four nodes in EAS, ECA, NAM, and China, in order to analyze the effect of additional cathode manufacturers on the market. The model includes 26 end-user nodes; nine of them for ESS and other electronics (one node for each region), four of them for EVs (EAS, ECA, NAM, and China), and four of them for non-LIB (EAS, ECA, NAM, and China). Finally, the recycling of LIB technologies is undergoing constant developments in order to improve the efficiency of the lithium recovery process. Specifically, the lithium recovery was started in 2019 with $1,000$ t LCE lithium compounds [64]. To this end, we consider nine recycling facility nodes in each region in order to analyze their effect on the lithium supply chain.

The cost function of lithium extraction is assumed to be linear with consideration of technological changes over the years. The technological change parameter, α_t^{T}, is obtained from Rosendahl and Rubiano [75] as 0.05%. Moreover, initial unit OPEX (c^{OPS}), unit CAPEX (c_g^{CA}), and initial annual extraction capacities for the existing and possible primary RMPs (δ_g^{P}) are obtained from Maxwell et al. [80]. This study also provides available resource amount which includes both discovered and undiscovered part of the deposit for 103 brine and hard rock deposits. However, by advancing in the exploration stage, the uncertainty of the available lithium amount in the deposit has been reduced since the reserve has been discovered. To this end, we assume the conversion rate from resource to reserve as 0.50 for brine deposit and 0.75 for hard rock deposit in order to find the reserve amount for each new and existing deposit.

The cost function of lithium recycling (c_f^{EFC}) is adapted from Rosendahl and Rubiano [75]. They provide a non-linear cost function which includes the efficiency of recycling facilities and technological changes, which is explained in the Appendix C. Specifically, they suggest that unit recovery cost is increased when the efficiency of the recycling facility increases since more energy is required. However, we consider a set of recycling facility efficiencies which includes the efficiency values ranging from 0.50 to 0.95 with 0.05 increments and let the model select the optimal efficiency among those values. In addition, as operational extraction costs, the recycling technologies advance over time which causes reduction in lithium recovery costs. Due to limited available data about lithium recovery from EOL batteries, we gather information about the recycling cost function parameters from Rosendahl and Rubiano [75]. Specifically, the lowest initial recycling unit cost (c^{OPS}), technological change parameter (α_t^{T}), and efficiency parameter (a_f^{EFN}) are obtained as $\$10,000$/t LCE, 0.05%, and 2, respectively.

Deterministic values are considered as the lithium requirement of each end-user sector. We obtain the total lithium requirement of each region in 2019 from Roskill [64] which also provides the global lithium requirement of three LIB sectors. Due to unavailable data for the regional lithium requirement of each sector, we multiply the share of regional EV sales by the global lithium requirement of EVs in order to find these requirements for each region. For example, in 2019, $1,059,529$ out of $2,015,738$ EVs (53% of global EV sale) are sold in China and global lithium requirement is $101,198$ t LCE. To this end, by assuming that every EV requires the same amount of lithium, we calculate the lithium requirement of EVs in China as $41,543$ t LCE. On the other hand, for the other two LIB sectors, we multiply the regional GDP share by the global lithium requirement of these sectors in order to obtain these regional requirements. After obtaining those for three LIB sectors, we subtract them from the regional lithium requirement in order to obtain the regional requirement for the non-LIB sector. Moreover, the growth rate of the requirement of each sector from 2020 to 2040 is provided by Benchmark Mineral Intelligence [65]. To the best of our knowledge, there are not any studies that provide regional growth rates for each sector. To this end, we assume that the lithium requirement of every region of a sector grows with the same rate in order to project regional lithium requirement for four sectors with respect to 2019 requirement as a base year.

We consider that the transportation of materials between two regions is done by cargo ships. The unit cost of material transportation is obtained as $/t LCE. Moreover, PAC, EAS, ECA, LAC, MEN, NAM, SAS, and SSA consist of 19, 18, 58, 42, 21, three, eighth, and 48 countries, respectively. To this end, we calculate the midpoint of a region by finding the center of gravity with respect to the GDP per capita of the countries which are located in the region. For example, if the region consists of two countries and the first country has higher GDP per capita, then the midpoint is near the first country. Note that we consider capital cities in order to represent countries' locations, and GDP per capita in 2020 are obtained from World BankThe World Bank [81]. Then, the distance between two regions is calculated with respect to the midpoints of these regions. Note that, since China is a single country, we consider Peking's latitude and longitude as the location of China.

In this study, we consider the GWP of raw material production, recycling, and transportation. To this end, we gather information on GWP for the production of lithium carbonate from brine and hard rock mine from Rosendahl and Rubiano [75] as 2.7 t CO_2/t LCE and 20.4 t CO_2/ t LCE, respectively. Note that these GWP values consider extraction operations, conversion process from lithium concentration to lithium carbonate form, and transportation. To this end, we do not consider GWP of transportation from hard rock mines to mineral conversion plants in order to avoid double counting. On the other hand, Mohr et al. [82] provide the GWP of lithium recovery from three cathode types of LIB namely, lithium nickel cobalt aluminum oxide, lithium nickel manganese cobalt oxide, and lithium iron phosphate as 0.10 t CO_2/ t LCE, 0.08 t CO_2/ t LCE, and 0.23 t CO_2/ t LCE, respectively. In this study, we do not consider different cathode types since the lithium requirements for each type which are 8.97 kg LCE, 11.31 kg LCE, and 13.80 kg LCE, respectively, are not significantly different. To this end, we take an average of these GWP values, which is 0.14 t CO_2/ t LCE as a GWP of lithium recovery from EOL battery. Finally, GWP for material transportation is obtained from GREET which is an open-source modeling tool proposed by Argonne National Laboratory as 4.59 t CO_2/ mile [83].

SCC is an estimated dollar value for the economic damages by emitting one additional ton of carbon dioxide into the atmosphere. Since carbon dioxide emissions have been affecting the economy over the years, the discount rate of the SCC is a critical parameter. In this study, we use two sets of carbon costs with 3% of discount rate that are proposed by the Interagency Working Group on the Social Cost of Greenhouse Gases (IWG), as shown in Table 4.1 [84]. Moreoever, we also consider the value of SCC as zero in order to analyze its effect on the lithium market. To this end, we have three scenarios regarding the values of SCC over the years from 2020 to 2050, namely, SCC 0, SCC 1 and SCC 2 as shown in Table 4.1.

Table 4.1 Carbon costs from 2020 to 2050 [84] for Three Scenarios

Emissions Year	2020	2025	2030	2035	2040	2045	2050
SCC 0	0	0	0	0	0	0	0
SCC 1	51	56	62	67	73	79	85
SCC 2	152	169	187	206	225	242	260

4.4.2 Computational Results and Discussions

In this section, we present our computational results obtained from the proposed model **CLLM** based on six different considerations, namely Model 1 through Model 6.

In 2019, cathode manufacturers are mainly concentrated in China and East Asia with 72% and 27% of global cathode powder production capacity, while the rest of the capacity is concentrated in NAM and EAS [64]. However, investors are planning to increase the concentration of these manufacturers in NAM and EAS. Specifically, the US government aims to increase domestic production and processing of critical raw materials in LIB and reduce its dependence on foreign supply. To this end, we consider two sets of cathode manufacturers, which consist of the manufacturers in China and EAS in Model 1 and in China, EAS, NAM, and ECA in Model 2, respectively, as shown in Table 4.2. For these models, SCC is not considered since it is not applied in the market. However, we also conduct sensitivity analysis based on two different scenarios of SCC, namely SCC 1 and SCC 2 as discussed in the previous section. In particular, Model 3 and Model 5 consider two cathode manufacturer nodes in China and EAS, while they consider SCC 1 and SCC 2 as environmental factors, respectively. On the other hand, Model 2 and Model 6 consider four cathode manufacturers in China, EAS, NAM, and ECA under SCC 1 and SCC 2 scenarios, respectively. Details for all six models are provided in Table 4.2. Note that the results discussed in this section do not represent the forecast for the lithium market. Instead, they provide reference points in order to analyze the effect of SCC and additional cathode manufacturers on the lithium market.

Note that our model **CLLM** includes 23,214 constraints, 23,029 variables, 2,276 of which are binary variables. In order to solve the problem, **CLLM** is coded in AMPL and solved using CPLEX 20.1.0.0. All computational runs are made on a Apple Mac Book Pro having an Intel Core i5-2600 CPU 1.40 GHz processor and 8 GB of RAM. Moreover, the optimal solution is provided for all Models 1-6 within 2 to 4 minutes.

Table 4.2 The considerations for location of cathode manufacturers and set of SCC in CLLM

Model	Location of Cathode Manufacturers	SCC
Model 1	China and EAS	SCC 0
Model 2	China, EAS, NAM, and ECA	SCC 0
Model 3	China and EAS	SCC 1
Model 4	China, EAS, NAM, and ECA	SCC 1
Model 5	China and EAS	SCC 2
Model 6	China, EAS, NAM, and ECA	SCC 2

4.4.2.1 Base Cases

In this section, we first analyze the closed-loop lithium supply chain by considering two sets of cathode manufacturers as given in Model 1 and Model 2. Specifically, these two models jointly consider the production of primary and secondary RMPs, annual production capacity of each primary RMP, their import and export decisions, EOL battery storage decisions of all supply chain components, as well as investment decisions on possible primary raw material deposits.

Table 4.3 Optimal solution of total discounted cost, cumulative GWP, and GWP caused by primary and secondary RPMs, transportation for Model 1 and Model 2

Result		Model 1	Model 2
Total Cost [5]	[M $]	157,083	155,687
Total GWP	[Mt CO_2]	11,654	7,318
GWP by Primary RMP	[Mt CO_2]	541	527
GWP by Secondary RMP	[Mt CO_2]	1.23	1.23
GWP by Transportation	[Mt CO_2]	11,110	6,789

Table Table 4.3 shows the optimal solutions for Model 1 and Model 2. Specifically, the optimal solution for Model 1 yields a total discounted cost[5] of $ 157,083 million (M). This cost covers OPEX of primary RMPs, secondary RMPs and mineral conversion plants, CAPEX of primary RMPs, and transportation costs throughout the supply chain. If additional cathode manufacturers in NAM and ECA are also taken into account, the optimal solution for Model 2 yields the total discounted cost of $ 155,687 M. Total GWP

[5] The total cost includes OPEX of primary and secondary RMPs, mineral conversion plants, CAPEX of primary RMPs, transportation cost, and SSC of transportation, lithium carbonate production and recovery.

is calculated based on the optimal decisions prescribed by the **CLLM** into three categories, primary RMPs' (brine and hard rock), and secondary RMPs' production and transportation. GWP for Model 1 by these categories are 542, 1.23, and 11,110 million (M) t CO_2, respectively. On the other hand, the corresponding GWP for Model 2 are 527, 1.23, and 6,789 M t CO_2, respectively.

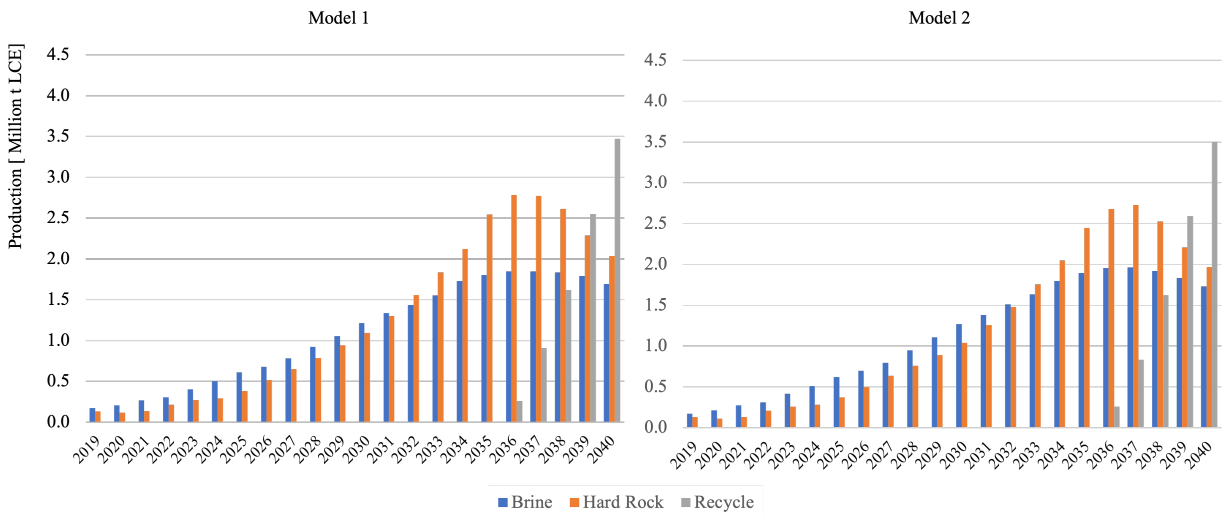


Figure 4.3 Optimal annual production of primary and secondary RMPs for Model 1 and Model 2

Figure 4.3 shows the optimal annual lithium productions by brine, hard rock, and recycling for Model 1 on the left and Model 2 on the right. In this figure, the production of these operations is highlighted in red, blue, and grey, respectively. The results of Model 1 show that brine operation is a dominant position in the raw material supply at the beginning of 20 years, whereas hard rock production is dominated after 2031. Moreover, the optimal solution suggests that the lithium recovery from EOL batteries starts in 2036 and has an exponential growth in the following years. On the other hand, in Model 2, the optimal brine production has a higher share of the market until 2032, then hard rock production has the highest share of the market among other lithium sources. Recycling operations start in 2036, becoming the main lithium source after 2039. Moreover, the total amount of recovered lithium from EOL LIB is 8,802 for Model 1 and Model 2.

We also calculate the Herfindahl-Hirschman Index (HHI) for countries based on the optimal annual production of primary RMPs in these countries for Model 1 and Model 2. HHI for Model 1 in 2019 is 4318, which reduces to 1459 by 2040. HHI for Model 2 is 4294 in 2019, and it decreases to 1429 by 2040. This downward trend is caused by the new production capacity of primary RMPs in countries where there is no primary RMP in 2019. For example, according to the optimal solution for Model 1, the new hard rock mine in Austria starts production in 2023, which decreases production shares of other countries in 2023. To this end, the HHI in 2031 is lesser than the HHI in 2032 for Model 1. Specifically, for Model 1, the annual primary raw material production is concentrated in Argentina, Australia, and Chile in 2019. Over time, the production in different countries increases their annual production. In 2040, the main production occurs in Argentina, Australia and Chile, whereas 39% of the lithium in the extraction stage is produced in Afghanistan, Austria, Brazil, Canada, China, Democratic Republic of Congo, Finland, Ghana, Mali, Mongolia, Namibia Portugal, Russia, Tibet, the US, Uzbekistan, and Zimbabwe. Similar to Model 1, in Model 2, the optimal annual production in 2019 is concentrated in Australia, Argentina, and Chile, while they produce 61% of primary raw material in 2040. The rest of the lithium is produced by the same set of countries in Model 1 except Mali which does not produce.

4.4.2.2 Sensitivity Analysis

In this section, we conduct sensitivity analysis in order to test the sensitivity of discounted cost to the changes in SCC and the location of cathode manufacturers. Regarding SCC, we consider two scenarios, namely SCC 1 and SCC 2 as provided in Table 4.1. Models 3-4 and Models 5-6 consider same scenarios for SCC (SCC 1 and SCC 2, respectively) while they consider different locations of cathode manufacturers. The optimal solutions for Models 3-6 are shown in Table 4.4. The results for Model 3 and Model 4 show that the total discounted costs for the next 20 years are \$ 441,807.19 M and \$ 278,098.63 M, respectively. Moreover, the optimal carbon emission for Model 3 is 6,189 M t CO_2, specifically, 642, 1.59, and 5,545 caused by primary RMP, secondary RMP, and material transportation, respectively. On the other hand, the optimal GWP for Model 4 is 2,275 M t CO_2. Specifically, primary RMP and secondary RMP operations and transportation cause GHG emissions of 165, 2.06, and 2,170 M t CO_2, respectively. In addition, the optimal total discounted costs for Model 5 and Model 6 are $\$ 978,874$ M and $\$ 477,256$ M, respectively. However, the optimal GWP for Models 5-6 is similar to those for Models 3-4, respectively.

Table 4.4 Optimal solution of total discounted cost, cumulative GWP, and GWP caused by primary and secondary RPMs, transportation for Model 3, Model 4, Model 5, and Model 6

Result		Model 3	Model 4	Model 5	Model 6
Total Cost [5]	[M \$]	441,807	278,099	978,874	477,256
Total GWP	[Mt CO_2]	6,189	2,275	6,184	2,266
GWP by Primary RMP	[Mt CO_2]	642	165	642	165
GWP by Secondary RMP	[Mt CO_2]	1.59	2.06	1.59	2.06
GWP by Transportation	[Mt CO_2]	5,545	2,108	5,540	2,099

Figure 4.4 shows the optimal annual lithium production by brine, hard rock, and recycling for Model 3 on the left and Model 4 on the right. In this figure, similarly to Figure 4.3, the production of these operations is highlighted in red, blue, and grey, respectively. The results of Model 3 show that hard rock

operation is mainly dominant in the raw material supply at the beginning of 20 years. Moreover, the optimal lithium recovery from EOL batteries starts in 2034, which is earlier than beginning of recycling operations in Model 1. On the other hand, according to Model 2, the optimal brine production is in the dominant position, while hard rock production has a small share of the raw material supply in the market. In addition, recycling operations start in 2031, becoming the primary lithium source after 2039.

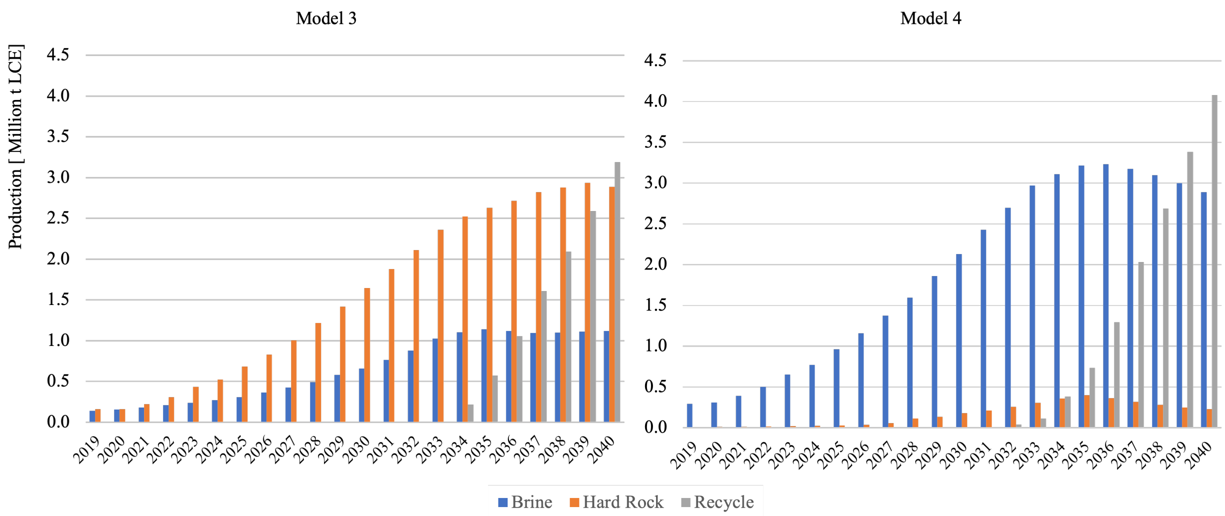


Figure 4.4 Optimal annual production of primary and secondary RMPs for Model 3 and Model 4

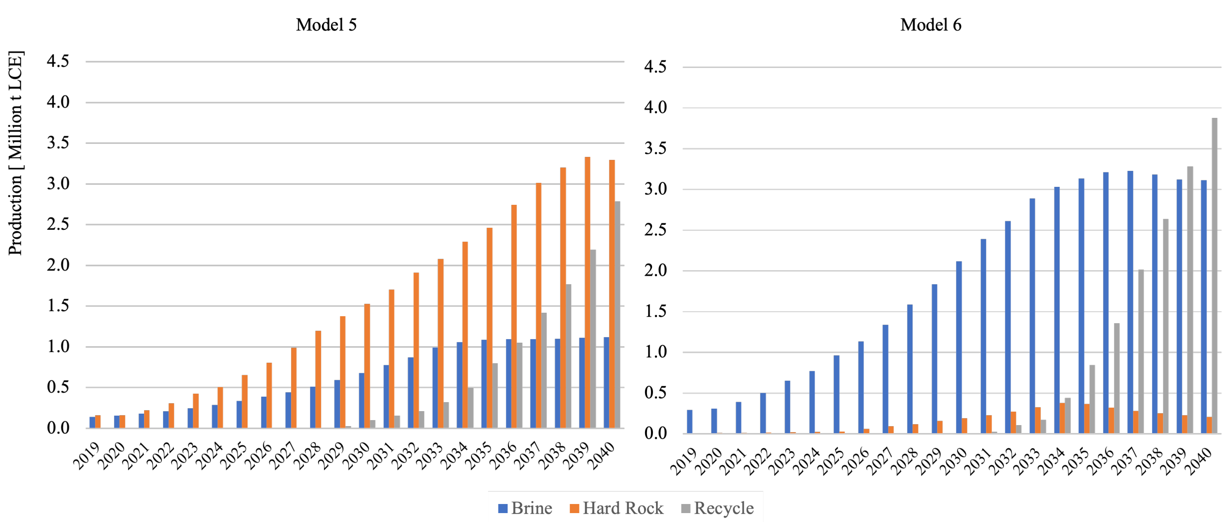


Figure 4.5 Optimal annual production of primary and secondary RMPs for Model 5 and Model 6

Figure 4.5 illustrates the optimal annual lithium production by brine, hard rock, and recycling for Model 5 on the left and Model 6 on the right. In this figure, the aforementioned highlights are used in order to represent brine, hard rock, and recycling facilities' annual productions. The annual production unit is Mt LCE. Similar to Model 3 results, the optimal results of Model 5 show that hard rock operation is mainly a dominant position in the raw material supply at the beginning of 20 years. Moreover, the optimal lithium recovery from EOL batteries starts in 2029. On the other hand, according to Model 6, the optimal brine production is dominant in the market. In addition, recycling operations are expected to start in 2031.

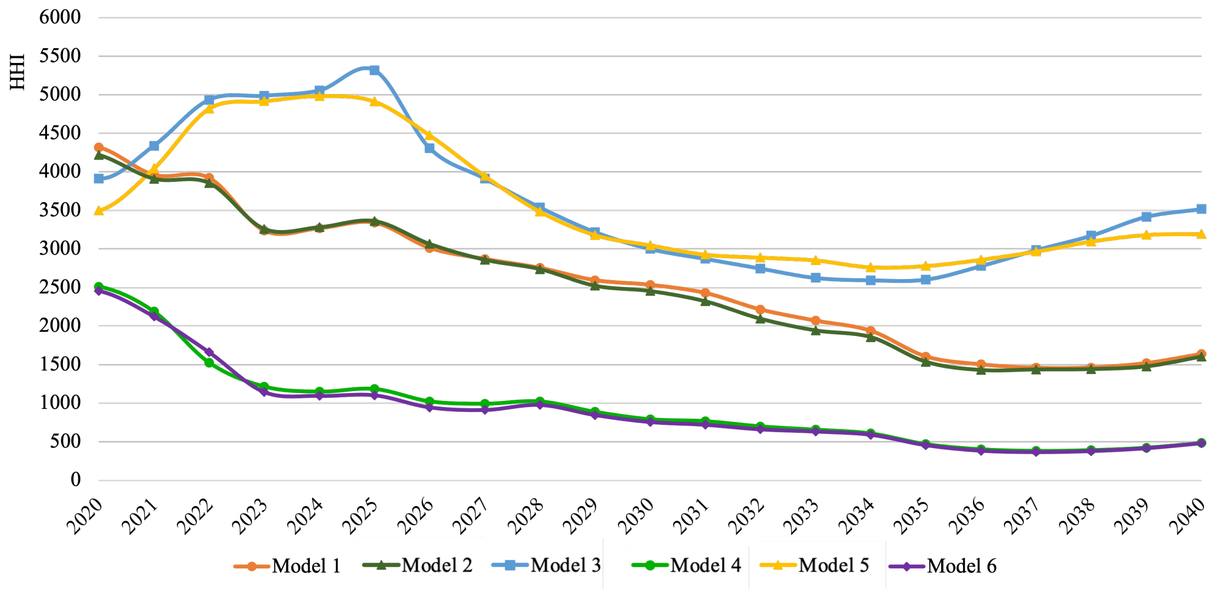


Figure 4.6 The HHI value of six models from 2019 to 2040.

In this study, we investigate the market concentration by using HHI under different SCC scenarios. Figure 4.6 illustrates the HHI for six models from 2019 to 2040 which are represented by orange, dark green, blue, green, yellow, and purple lines, respectively. As shown in the figure, the optimal HHI trends of Model 3 and Model 4 over the next 20 years are similar to those of Model 5 and Model 6, respectively, since the optimal annual productions of countries are not significantly different in Models 3-5 and Models 4-6. The variations in trends are caused by starting year of recycling operations. According to the optimal solution of Model 3, the HHI increases from 3,915 to 5,316 between 2019 and 2025. It decreases to 2,495 by 2034 and is stable by 2040. On the other hand, the optimal HHI of Model 4 decreases from 2,547 to 500 by 2040.

4.4.2.3 Discussion

In this study, we discuss six models under different scenarios of SCC and cathode manufacturer locations in order to analyze their effects on the total discounted cost, annual production of primary and secondary RMPs, GWP of these productions, material transportation, and country-level market concentration for the extraction stage. First, we compare models that consider the same set of cathode manufacturers; then, we compare the models that consider the same scenarios of SCC.

The comparison of the optimal solutions for Model 1 and Model 2 reveals that additional cathode manufacturers in NAM and ECA do not affect the total discounted cost since the unit transportation cost is relatively smaller than the other costs. By contrast, they cause a 61% of reduction in the GWP by transportation over 20 years, while it does not affect GWP by secondary RMP. Moreover, they increase annual brine production and decreases annual hard rock production since most brine deposits concentrate in LAC while not affecting annual lithium recovery from EOL LIB. Since these changes are not significant, the market concentrations of countries with respect to annual primary RMPs' production for both models are not significantly different. Specifically, the optimal solutions of these models show that lithium is produced in Australia, Argentina, and Chile in the extraction stage at the beginning. In 2040, nearly 39% of the required lithium is produced in different countries. To this end, country-level market concentration in the extraction stage decreases from 2019 to 2040.

The comparison of the optimal solutions for Model 3 and Model 4, under the scenario of SCC 1, indicates that additional cathode manufacturers in NAM and ECA decrease the total discounted cost by 63%; in addition, they cause a 37% reduction in the GWP over 20 years. Specifically, they decreases primary RMP and transportation GWP by 26% and 38%, respectively. However, it increases secondary RMP's GWP by 30% since the cumulative recovery of lithium is increased from 11,322 to 14,746 Mt LCE. According to the optimal solution for Model 3, although lithium carbonate production in mineral conversion plants that originated from hard rock deposits has a higher GWP than lithium carbonate conversion of brine sourced lithium, hard rock production is advantageous. This is because most of the brine deposits are located in LAC which requires the higher distance transportation to cathode manufacturers. However, additional cathode manufacturers in NAM and ECA prioritize the brine operations since they reduce the required transportation to cathode manufacturers. Moreover, they also reduce the country-level market concentration. Specifically, the main primary raw material productions occur in Australia without considering additional cathode manufacturers, while those in NAM and ECA are replaced by Chile.

As discussed in Section 4.4, the objective function of the proposed model includes both are direct cost and GWP. In order to consider these two factors jointly, we use two scenarios of SCC as the monetization of GWP since it plays a critical role in prioritizing GWP over the direct cost. According to optimal solutions for Models 5-6, increasing SCC values decreases the cumulative GWP. However, the marginal GWP is reduced over time since it gets closer to the minimum GWP. To this end, the optimal solution for Models 3-4, which consider the scenario of SCC 1, are not significantly different from those for Models 5-6, which consider the scenario of SCC 2. Therefore, increasing SCC can cause an increase in total discounted costs without a significant environmental effect.

We investigate the effect of SCC on the lithium market with respect to the annual production of primary and secondary RMPs, GWP, and the country-level market concentration, considering two cathode manufacturers in China and EAS. The comparison of optimal annual production of primary and secondary RMPs for Models 1, 3, and 5 implies that SCC increases the market share of hard rock operations since GWP caused by material transportation prioritizes the deposits near China or East Asia, which are hard rock deposits in Australia, China, and Russia. Moreover, increasing of SCC causes reset to an earlier year of recycling operations. For example, the recycling operations start in 2036, 2034, and 2029, respectively. Although considering SCC increases the cumulative lithium recovery from EOL batteries, the cumulative lithium recoveries are the same for Model 3 and Model 5. The comparison of the optimal GWP for these models suggests that the consideration of SCC reduces GWP by 53%. Specifically, GWP caused by material transportation is reduced by 50%, while GWP caused by primary and secondary RMPs are increased by 19% and 29%, respectively. Finally, according to these three models, consideration of SCC forces the model to prioritize the deposit near China and East Asia, where the cathode manufacturers are located, in order to reduce GWP. Since annual capacity growth is limited by 20%, existing hard rock deposits meet the main lithium requirement from 2019 to 2025. Then, the annual capacities of new deposits in China, Russia, and Australia are increased. To this end, the HHI has a peak point in 2025, then country-level market concentration reduces next ten years.

Considering four cathode manufacturers in China, EAS, NAM, and ECA, we investigate the effect of SCC on the lithium supply chain with respect to the annual production of primary and secondary RMPs, GWP, and the country-level market concentration. The optimal annual production for Model 2, Model 4, and Model 6 indicate that consideration of SCC prioritizes brine deposits over hard rock deposits since they have low GWP than hard rock operation. Moreover, it resets to an earlier year of recycling operations, specifically, from 2036 to 2032. However, increasing SCC does not significantly affect the share of these operations since GWP is at the minimum point under the scenario of SCC 1. It also does not change the starting year of the recycling operation. The optimal GWPs for these three models reveal that the consideration of SCC decreases 31% of total GWP, specifically, 31% reduction on GWP caused by primary RMPs and material transportation. On the other hand, it increases GWP caused by secondary RMPs since it increases the cumulative recovery of lithium from EOL LIB. However, the additional SCC does not decrease the total GWP, which means that marginal SCC on the cumulative GWP is equal to zero. To this end, it also does not affect the country-level market concentration in the extraction stage. However, consideration of SCC decreases the HHI over the next 20 years.

4.5 Conclusion

This paper examines the closed-loop lithium supply chain as a network model that includes primary and secondary RMPs, mineral conversion plants, cathode manufacturers, battery and final good manufacturers, and end-users from four sectors. In this study, we assume all lithium in the extraction stage and recovered lithium are transformed into lithium carbonate. In addition, all end-users require lithium as the chemical form of lithium carbonate. To address this problem, a mixed-integer linear programming formulation is proposed that prescribes optimal annual production decisions of primary and secondary RMPs, mineral conversion plants, and cathode manufacturers, annual capacity decisions of primary RMPs, recovery efficiency decisions of secondary RMPs, storage decisions of EOL LIB, and investment decisions of new primary RMPs in order to minimize the total discounted cost. This cost includes OPEX for lithium carbonate production originating from brine, hard rock, or recycling facilities, CAPEX for brine and hard rock deposits, transportation cost of the supply chain, and monetized GWP for lithium carbonate production and transportation. We also consider SCC in order to minimize the GWP of lithium production for the required lithium of the future. To the best of our knowledge, this is the first study that jointly considers all these decisions for the lithium carbonate production in the global lithium market in order to examine the effect of SCC and cathode manufacturers' location.

In this study, we provide six versions of the problem, including different scenarios regarding the locations of cathode manufacturers and SCC values. Our results suggest that additional cathode manufacturers in NAM and ECA can reduce the required material transportation distance. Without consideration of SCC, it does not significantly affect total discounted cost since transportation cost is relatively smaller than other cost parameters. However, it reduces the GWP by increasing production with respect to the required distance from deposits to cathode manufacturers. Specifically, it increases the brine production in Argentina and hard rock production in Mongolia, Russia, and Uzbekistan. At the same time, it decreases the hard rock production in Australia, Afghanistan, Mali, and the Democratic Republic of Congo. Comparing the optimal solutions of these models reveals that the SCC can reduce the GWP of lithium carbonate productions, especially GWP caused by material transportation. Moreover, SCC prioritizes hard rock operations rather than brine operations with two cathode manufacturer nodes in EAS and China while prioritizing brine operations rather than hard rock operations with additional cathode manufacturer nodes in NAM and ECA. However, without consideration of SCC, the share of brine and hard rock operation of annual lithium production in the extraction stage is closed to each other. In addition, considering the scenario of SCC 2 rather than SCC 1 does not significantly affect GWP and country-level market concentration while increasing total discounted cost since SCC 2 is greater than SCC

1. Moreover, the effect of SCC on the country-level market concentration depends on the cathode manufacturers in the market. Specifically, when the cathode manufacturers' concentration is reduced with additional cathode manufacturers in NAM and ECA, the country-level market concentration of primary RMPs is also reduced. However, because of the production capacity constraints of the deposits, the country-level market concentration increases during the first five years. Specifically, the existing deposits in Australia, which are closer to mineral conversion plants in China, supply the lithium requirements in this period. Then, the deposits which are also closer to China but not located in Australia increase their production over time. Specifically, the optimal solution yields that the hard rock deposits in Russia are the main primary RMP in 2040.

In this study, the lithium requirements for each region play a critical role in the model decisions. Due to the limited availability of the data at the region level, we make several assumptions in order to obtain the amount of lithium requirement for four sectors in nine regions. To this end, we aim to estimate these requirements to be used in our model as a future study. In addition, we consider that all lithium chemical form is lithium carbonate in the market in order to obtain reference points for the effects of additional cathode manufacturers and SCC. However, in 2019, 25% of the lithium is used in the lithium hydroxide form [64]. Since the brine operations are advantageous for lithium carbonate production, the proposed model prioritizes brine operations rather than hard rock operations. As a future study, the lithium hydroxide can be added to the model in order to obtain accurate decisions; then, the model can be used to forecast the market.

CHAPTER 5

CONCLUSION

Our research considers the investment decisions of policymakers in the energy and mineral supply chain with the objective of maximizing social welfare or minimizing cost. Mineral and energy markets need to be analyzed both globally and regionally since both markets are directly related to clean energy transitions. We investigate different markets example, EU natural gas and closed-loop lithium markets.

In Chapter 2, we estimate the residential natural gas demand for 23 European countries. We use three different causality estimation versions of traditional IV and post-LASSO IV methodologies. To the best of our knowledge, this is the first initiative that estimates the causality of natural gas prices for individual 23 European countries as well as for the full panel of these countries. We also compare these methodologies for natural gas price estimates in order to investigate other countries' weather effects on European countries' natural gas consumption. The methods are performed by three criteria, namely, reasonableness, relevance, and validity. The results reveal that at the country level, two-stage post-LASSO IV outperforms by satisfying the greatest number of these three criteria for the greatest number of countries, while two-stage post-LASSO IV has the most reliable results at the panel level. Moreover, our preferred estimates suggest that country-level price elasticities range from -0.98 to -0.09, with a median of -0.58, in line with estimates from previous literature.

In Chapter 3, we address the EU natural gas supply chain problem where possible new infrastructural projects may be invested in order to maximize the social welfare of consumers in EU countries. We propose a mixed-integer non-linear programming model, namely, EGIM, which may help policymakers who are interested in the investment of possible infrastructure projects with a limited budget. The EGIM provides optimal amounts of extracted natural gas for producers, transported gas via pipeline or LNG terminals, traded gas between supply chain actors, stored gas for mid-streamers, consumed for three sectors (i.e., household, industry, and power), as well as new pipeline or LNG regassification terminal investments. We employ the EGIM under four scenarios which are the combination of considering investment decisions and Russian gas interruptions. Our results reveal that new infrastructural investments cause the higher welfare, nearly 3 billion, than the market without these investments, as the consumer welfare theory also suggests. However, Our results show that the production stage of the market without these investments is more competitive than the stage of the market with these investments according to their country-level HHI, calculated as 1700 and 1719, respectively. In addition, the results of monthly natural gas consumption and price indicate that making the investment in new infrastructural project increase the social welfare of consumers in Bulgaria, Denmark, Croatia, and Sweden, while reducing it in Lithuania. Our results also suggest that Russian gas interruption causes reduced social welfare and cumulative available gas in the EU natural gas market by 10% and 15%, respectively. Specifically, the social welfare of consumers in Finland, Latvia, and Estonia are significantly affected by Russian gas interruptions. However, Estonia can reduce these interruptions by investing in the reggasfication terminal.

In Chapter 4, we propose a mixed-integer linear programming model, namely, CLLM, for the global closed-loop lithium supply chain. We find that the minimum cost of extraction, recycling, and transportation in order to meet the next 20 years of global lithium requirement is $157,083 M where the cathode manufacturers are located in China, Japan, and South Korea. These results reveal similar market shares for brine and hard rock suppliers until 2033. In 2036, hard rock production peaks when lithium recovery from EOL LIBs starts. Then, recycling facilities are the dominant producers in the market after 2039. The GWP is calculated based on the optimal decisions prescribed by the proposed model as $11,654$ Mt CO_2, which includes 541, 1.23, and $11,110$ Mt GWP for primary, secondary RMPs, and transportation, respectively. Moreover, additional cathode manufacturers in the US and Europe do not significantly affect the total cost since the transportation cost is relatively small than the other cost parameters. However, it reduces the GWP to $7,318$ Mt CO_2. We also investigate the effect of SCC on the market. According to our results, SCC prioritizes one of the supply sources depending on the cathode manufacturer's location. When the cathode manufacturers are located in China and EAS, brine operations are more expensive than hard rock operations because of material transportation. Otherwise, hard rock operations are more expensive than brine since lithium carbonate production from hard rock-sourced lithium has a higher GWP. Moreover, SCC can reduce the GWP from $11,654$ to $6,189$ Mt CO_2 without additional cathode manufacturers and from $7,318$ to $2,275$ Mt CO_2 with additional cathode manufacturers. Finally, the country-level market concentration for the extraction stage does not significantly change by the additional cathode manufacturers. The results of the models imply that decreasing market concentration in one of the supply chain stages may decrease the market concentration in other stages. Specifically, decreasing market concentration in cathode manufacturers decrease the country-level HHI in the extraction stage from $4,000$ to $2,500$.

REFERENCES

[1] British Petroleum Company. BP statistical review of world energy. Technical report, Barcelona: British Petroleum Co., 2021.

[2] Ben Holland, Scott Johnson, Jamie Rush, Anna Wong, and Tom Orlik. How war in ukraine threatens the world's economic recovery. https://www.bloomberg.com/news/features/2022-02-25/war-in-ukraine-how-the-ukraine-russia-conflict-could-impact-the-global-economy, 2022.

[3] The European Commission. Questions and answers on reducing methane emissions in the energy sector. https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_6684, 2021.

[4] C Gürsan and V de Gooyert. The systemic impact of a transition fuel: Does natural gas help or hinder the energy transition? *Renewable and Sustainable Energy Reviews*, 138:110552, 2021.

[5] Katrin Rehdanz. Determinants of residential space heating expenditures in Germany. *Energy Economics*, 29(2):167–182, 2007.

[6] Peter H.G. Berkhout, Ada Ferrer-i Carbonell, and Jos C. Muskens. The ex post impact of an energy tax on household energy demand. *Energy Economics*, 26(3):297–317, 2004.

[7] Helena Meier and Katrin Rehdanz. Determinants of residential space heating expenditures in great britain. *Energy Economics*, 32(5):949–959, 2010.

[8] R A Rohde and Z Hausfather. The Berkeley earth land/ocean temperature record. *Earth System Science Data*, 12(4):3469–3479, 2020.

[9] Catherine Hausman and Ryan Kellogg. Welfare and distributional implications of shale gas. NBER Working Paper 21115, National Bureau of Economic Research, Cambridge, MA, 2015.

[10] Lucas W. Davis and Erich Muehlegger. Do Americans consume too little natural gas? An empirical test of marginal cost pricing. *RAND Journal of Economics*, 41(4):791–810, 2010.

[11] A Belloni, D Chen, V Chernozhukov, and C Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429, 2012.

[12] Hyunseung Kang, Anru Zhang, T Tony Cai, and Dylan S Small. Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. *Journal of the American Statistical Association*, 111(513):132–144, 2016.

[13] Frank Asche, Odd Bjarte Nilsen, and Ragnar Tveteras. Natural gas demand in the european household sector. *The Energy Journal*, 29(3):27–46, 2008.

[14] Gangadharrao S Maddala, Robert P Trost, Hongyi Li, and Frederick Joutz. Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators. *Journal of Business & Economic Statistics*, 15(1):90–100, 1997.

[15] Gang Liu. Estimating energy demand elasticities for oecd countries. a dynamic panel data approach. Discussion Papers No. 373, Research Department, Statistics Norway, 2004.

[16] Paul J. Burke and Hewen Yang. The price and income elasticities of natural gas demand: International evidence. *Energy Economics*, 59:466–474, 2016.

[17] O Dilaver, Z Dilaver, and Lester C. Hunt. What drives natural gas consumption in europe? analysis and projections. *Journal of Natural Gas Science and Engineering*, 19:125–136, 2014.

[18] Ronald Bernstein and Reinhard Madlener. Residential natural gas demand elasticities in oecd countries: an ardl bounds testing approach. FCN working paper 15/2021, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, 2011.

[19] Massimo Filippini and Nilkanth Kumar. Gas demand in the Swiss household sector. *Applied Economics Letters*, 28(5):359–364, 2021.

[20] Ioannis Kostakis, Sarantis Lolos, and Eleni Sardianou. Residential natural gas demand: Assessing the evidence from greece using pseudo-panels, 2012–2019. *Energy Economics*, 99:105301, 2021.

[21] Maximilian Auffhammer and Edward Rubin. Natural gas price elasticities and optimal cost recovery under consumer heterogeneity: Evidence from 300 million natural gas bills. NBER Working Paper 24295, National Bureau of Economic Research, Cambridge, MA, 2018.

[22] Michael J. Roberts and Wolfram Schlenker. Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. *American Economic Review*, 103(6): 2265–2295, 2013.

[23] Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Lasso methods for gaussian instrumental variables models. MIT Working Paper 11-14, Department of Economics, Massachusetts Institute of Technology, 2011.

[24] Victor Chernozhukov, Christian Hansen, and Martin Spindler. Post-selection and post-regularization inference in linear models with many controls and instruments. *American Economic Review*, 105(5): 486–90, 2015.

[25] Vira Semenova, Matt Goldman, Victor Chernozhukov, and Matt Taddy. Estimation and inference on heterogeneous treatment effects in high-dimensional dynamic panels. *arXiv preprint arXiv:1712.09988*, 2017.

[26] Wei Lin, Rui Feng, and Hongzhe Li. Regularization methods for high-dimensional instrumental variables regression with an application to genetical genomics. *Journal of the American Statistical Association*, 110(509):270–288, 2015.

[27] Frank Windmeijer, Helmut Farbmacher, Neil Davies, and George Davey Smith. On the use of the lasso for instrumental variables estimation with some invalid instruments. *Journal of the American Statistical Association*, 114(527):1339–1350, 2019.

[28] Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650, 2013.

[29] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[30] Alexandre Belloni and Victor Chernozhukov. Least squares after model selection in high-dimensional sparse models. *Bernoulli*, 19(2):521–547, 2013.

[31] Merve Olmez Turan and Tulay Flamand. Optimizing investment and transportation decisions for the European natural gas supply chain. Unpublished manuscript, Department of Economics and Business, Colorado School of Mines, 2022.

[32] IGU. Global gas report 2019. Technical report, International Gas Union, London, 2019.

[33] IEA. Market report series: Gas 2019 analysis. Technical report, International Energy Agency, Paris, 2019.

[34] Alla Khalitova. Energy trade brinkmanship between the european union and russia. Technical Report 1, 2011.

[35] EuroStat. European Union imports of energy products recent developments, 2019.

[36] Ole Gunnar Austvik. The Energy Union and security-of-gas supply. *Energy Policy*, 96:372–382, 2016.

[37] Rudolf G. Egging and Steven A. Gabriel. Examining market power in the european natural gas market. *Energy Policy*, 24:2762–2778, 2006.

[38] Ruud Egging, Steven A. Gabriel, Franziska Holz, and Jifang Zhuang. A complementarity model for the European natural gas market. *Energy Policy*, 36(7):2385–2414, 2008.

[39] E Allevi, L Boffino, Giuli M. E. De, and G Oggioni. Evaluating the impacts of the external supply risk in a natural gas supply chain : the case of the Italian market. *Journal of Global Optimization*, 70(2): 347–384, 2018.

[40] F. A. M. Boots, M. G.and Rijkers and B. F. Hobbs. Trading in the Downstream European Gas Market: A Successive Oligopoly Approach. *The Energy Journal*, 25(3):73–102, 2004.

[41] Peter Hartley and Kenneth B Medlock. The baker institute world gas trade model. http://www.bakerinstitute.org, 2005.

[42] Ruud Egging, Franziska Holz, and Steven A. Gabriel. The World Gas Model. A multi-period mixed complementarity model for the global natural gas market. *Energy*, 35(10):4016–4029, 2010.

[43] Daniel De Wolf and Yves Smeers. A stochastic version of a Stackelberg-Nash-Cournot equilibrium model. *Management Science*, 43(2):190–197, 1997.

[44] Gijsbert Zwart and Machiel Mulder. NATGAS: a model of the European natural gas market. Technical report, CPB Memorandum, 2006.

[45] Franziska Holz, Philipp M. Richter, and Ruud Egging. The role of natural gas in a low-carbon Europe: Infrastructure and supply security. *Energy Journal*, 37:33–59, 2016.

[46] R. Egging. Benders decomposition for multi-stage stochastic mixed complementary problems-applied to a global natural gas market model. *European Journal of Operational Research*, 226(2):341–353, 2013.

[47] Geir Hasle, Asgeir Tomasgard, Frode Romo, Marte Fodstad, and Kjetil Midthun. Optimization models for the natural gas value chain. In *Geometric Modelling, Numerical Simulation, and Optimization: Applied Mathematics at SINTEF*, pages 521–558. Springer Berlin Heidelberg, 2007.

[48] Daniel J.G. Crow, Sara Giarola, and Adam D. Hawkes. A dynamic model of global natural gas supply. *Applied Energy*, 218:452–469, 2018.

[49] Stevie Lochran. Gnome: A dynamic dispatch and investment optimisation model of the european natural gas network and its suppliers. *Operational Resource Forum*, 2(67), 2021.

[50] J P Deane, M Ó Ciaráin, and B P Ó Gallachóir. An integrated gas and electricity model of the EU energy system to examine supply interruptions. *Applied Energy*, 193:479–490, 2017.

[51] Franziska Holz, Christian von Hirschhausen, and Claudia Kemfert. A strategic model of European gas supply (GASMOD). *Energy Economics*, 30(3):766–788, 2008.

[52] Marte Fodstad, Ruud Egging, Kjetil Midthun, and Asgeir Tomasgard. Stochastic modeling of natural gas infrastructure development in Europe under demand uncertainty. *Energy Journal*, 37:5–32, 2016.

[53] András Kiss, Adrienn Selei, and Borbála Takácsné Tóth. A top-down approach to evaluating cross-border natural gas infrastructure projects in Europe. *Energy Journal*, 37(3):61–79, 2016.

[54] BP. BP statistical review of world energy, 2020.

[55] GIE. Storage database. https://www.gie.eu/index.php/gie-publications/databases/storage-database, 2019.

[56] ENTSO-G. Transmission capacity and system development maps. https://www.entsog.eu/sites/default/files/2020-01/ENTSOG_GIE_SYSDEV_2018-2019_1600x1200_FULL_063_clean.pdf, 2019.

[57] Peter Kotek, Pedro Crespo Del Granado, Ruud Egging, and Borbala T. Toth. European natural gas infrastructure in the energy transition. In *International Conference on the European Energy Market, EEM*, pages 1–6. IEEE, 2019.

[58] Robert Baron, Paul Bernstein, W. David Montgomery, and Sugandha D. Tuladhar. Updated macroeconomic impacts of LNG exports from the United States. Technical report, NERA Economic Consulting, Washington, DC, 2014.

[59] sea distance.org. Sea distance and port distance - online tool for calculation distances between sea ports. https://www.sea-distances.org/, 2019.

[60] Merve Olmez Turan, Ben Gilbert, and Tulay Flamand. How good are weather shocks for identifying energy elasticities? A LASSO-IV approach to European natural gas demand. Unpublished manuscript, Department of Economics and Business, Colorado School of Mines, 2022.

[61] IEA. Natural gas information. Technical report, International Energy Agency, Paris, 2019.

[62] Anne Neumann, Sophia Rüster, and Christian von Hirschhausen. Long-term contracts in the natural gas industry: Literature survey and data on 426 contracts (1965-2014), 2015.

[63] DOE. Critical minerals and materials: U.S. Department of Energy's strategy to support domestic critical mineral and material supply chains (FY 2021-FY 2031). Technical report, US Department of Energy, 2021.

[64] Roskill. Lithium: Outlook to 2030 seventeenth edition. Technical report, Londan, 2020.

[65] Benchmark Mineral Intelligence. Lithium forecast – 2nd quarter 2020. Technical report, Londan, 2020.

[66] Alexandru Sonoc, Jack Jeswiet, and Vi Kie Soo. Opportunities to improve recycling of automotive lithium ion batteries. *Procedia CIRP*, 29:752–757, 2015.

[67] Laurence Kavanagh, Jerome Keohane, Guiomar Garcia Cabellos, Andrew Lloyd, and John Cleary. Global lithium sources-industrial use and future in the electric vehicle industry: A review. *Resources*, 7(3), 2018.

[68] Basudev Swain. Recovery and recycling of lithium: A review. *Separation and Purification Technology*, 172:388–403, 2017.

[69] IEA. Tracking transport 2021. Technical report, Paris, 2021.

[70] Troy R Hawkins, Bhawna Singh, Guillaume Majeau-Bettez, and Anders Hammer Strømman. Comparative environmental life cycle assessment of conventional and electric vehicles. *Journal of Industrial Ecology*, 17(1):53–64, 2013.

[71] Maarten Messagie. Life cycle analysis of the climate impact of electric vehicles. *Journal of Life Cycle Assessment*, 2014:14, 2014.

[72] Lin Li, Fadwa Dababneh, and Jing Zhao. Cost-effective supply chain for electric vehicle battery remanufacturing. *Applied Energy*, 226:277–286, 2018.

[73] Sam Jaffe. Vulnerable links in the lithium-ion battery supply chain. *Joule*, 1(2):225–228, 2017.

[74] Elsa A Olivetti, Gerbrand Ceder, Gabrielle G Gaustad, and Xinkai Fu. Lithium-ion battery supply chain considerations: analysis of potential bottlenecks in critical metals. *Joule*, 1(2):229–243, 2017.

[75] Knut Einar Rosendahl and Diana Roa Rubiano. How effective is lithium recycling as a remedy for resource scarcity? *Environmental and Resource Economics*, 74(3):985–1010, 2019.

[76] Harald Ulrik Sverdrup, Kristin Vala Ragnarsdottir, and Deniz Koca. Integrated modelling of the global cobalt extraction, supply, price and depletion of extractable resources using the WORLD6 model. *BioPhysical Economics and Resource Quality*, 2(1):1–29, 2017.

[77] Xin Sun, Han Hao, Fuquan Zhao, and Zongwei Liu. Tracing global lithium flow: A trade-linked material flow analysis. *Resources, Conservation and Recycling*, 2017.

[78] Xin Sun, Han Hao, Philipp Hartmann, Zongwei Liu, and Fuquan Zhao. Supply risks of lithium-ion battery materials: An entire supply chain estimation. *Materials Today Energy*, 14:100347, 2019.

[79] Lin Li, Fadwa Dababneh, and Jing Zhao. Cost-effective supply chain for electric vehicle battery remanufacturing. *Applied Energy*, 226:277–286, 2018.

[80] Fleming Maxwell, Kannan Sangita Gayatri, and Eggert Roderick. Assessment of long-run availability of lithium: A cumulative availability curve. Critical Materials Institute Winter Meeting, 2022.

[81] World BankThe World Bank. Gdp per capita, 2022. data retrieved from World Development Indicators, https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

[82] Marit Mohr, Jens F Peters, Manuel Baumann, and Marcel Weil. Toward a cell-chemistry specific life cycle assessment of lithium-ion battery recycling processes. *Journal of Industrial Ecology*, 24(6): 1310–1322, 2020.

[83] Michael Q Wang. Development and use of greet 1.6 fuel-cycle model for transportation fuels and vehicle technologies. Technical report, Argonne National Lab., IL (US), 2001.

[84] IWG. Technical support document: social cost of carbon, methane, and nitrous oxide interim estimates under executive order 13990. Technical report, Washington, D.C., 2021.

APPENDIX A

SUPPLEMENTARY INFORMATION FOR CHAPTER 2

Table A.1 provides the results of four IV, post-LASSO IV, and two-stage post-LASSO IV models for each 23 EU country. The results includes price elasticity, standard error of price elasticity, First-stage F-statistics, and Sargan statistic p-value. In the table, abbreviation of "PL", "2S PL", and "UR" represent post-LASSO and two-stage post-LASSO.

Table A.1 Results of four IV models, two-stage post-LASSO IV, and post-LASSO IV models at country-level

Country	Model	Price Elasticity	Std. Error of Price Elasticity	First-stage F-statistics	Sargan Test (p-Value)
Austria	Model IV 1	−1.18	0.95	0.57	0.96
	Model IV 2	1.71	2.02	2.30	0.99
	Model IV 3	0.14	0.50	2.34	0.95
	Model IV 4	0.17	0.48	13.15	0.92
	PL IV	−0.37	0.16	2.74	≤ 0.05
	2S PL IV	−0.67	0.25	3.87	≤ 0.05
Belgium	Model IV 1	0.26	0.41	1.44	1.00
	Model IV 2	0.09	0.35	2.86	0.92
	Model IV 3	0.26	0.20	2.39	0.50
	Model IV 4	0.32	0.18	4.46	0.26
	PL IV	−0.64	0.26	3.33	≤ 0.05
	2S PL IV	−0.16	0.08	2.76	0.21
Bulgaria	Model IV 1	−1.32	2.02	0.63	0.78
	Model IV 2	−1.01	0.58	3.18	0.75
	Model IV 3	−1.96	0.81	2.63	0.89
	Model IV 4	−0.83	0.29	6.77	0.39
	PL IV	−0.49	0.71	1.31	≤ 0.05
	2S PL IV	0.06	0.53	3.52	0.70
Croatia	Model IV 1	0.64	0.51	0.72	0.95
	Model IV 2	0.10	0.42	0.72	1.00
	Model IV 3	0.01	0.22	0.73	0.03
	Model IV 4	−0.04	0.20	2.16	0.04
	PL IV	−0.75	0.32	1.80	0.29
	2S PL IV	−0.10	0.13	2.79	0.78
Czech Republic	Model IV 1	2.38	1.42	3.46	0.51
	Model IV 2	−0.21	3.07	3.02	0.47
	Model IV 3	1.31	0.55	4.45	0.41
	Model IV 4	2.45	0.72	7.40	0.12
	PL IV	−0.58	0.18	4.05	0.06
	2S PL IV	1.00	0.25	8.71	0.12
Denmark	Model IV 1	−0.63	0.52	1.04	0.57
	Model IV 2	0.58	0.76	0.44	0.73
	Model IV 3	−0.58	0.30	2.66	0.82

Table A.1 Continued.

Country	Model	Price Elasticity	Std. Error of Price Elasticity	First-stage F-statistics	Sargan Test (p-Value)
Denmark	Model IV 4	−0.33	0.42	0.98	0.99
	PL IV	−0.11	0.30	2.70	0.03
	2S PL IV	−0.30	0.14	2.38	0.23
Estonia	Model IV 1	0.20	0.44	0.21	0.38
	Model IV 2	−0.03	0.26	0.76	0.90
	Model IV 3	0.22	0.12	1.59	0.17
	Model IV 4	0.22	0.12	1.90	0.75
	PL IV	−0.18	0.09	2.96	0.23
	2S PL IV	−0.04	0.06	2.40	0.04
France	Model IV 1	−0.99	0.81	1.02	0.88
	Model IV 2	0.19	0.93	1.15	1.00
	Model IV 3	−0.92	0.28	6.06	0.40
	Model IV 4	−1.33	0.38	9.17	0.29
	PL IV	−0.22	0.47	0.81	0.32
	2S PL IV	−0.35	0.10	1.92	0.34
Germany	Model IV 1	−1.19	1.22	1.07	0.73
	Model IV 2	−1.12	1.42	1.94	0.78
	Model IV 3	−1.25	0.51	2.42	0.58
	Model IV 4	−0.86	0.77	3.36	0.56
	PL IV	−1.17	0.32	3.01	0.41
	2S PL IV	−0.60	0.30	2.62	0.25
Greece	Model IV 1	−2.90	3.25	0.56	0.57
	Model IV 2	−1.44	2.76	0.22	0.60
	Model IV 3	−2.46	0.73	5.73	0.45
	Model IV 4	−2.08	0.64	5.46	0.48
	PL IV	−0.23	0.13	5.22	0.03
	2S PL IV	−0.98	0.23	2.58	0.25
Hungary	Model IV 1	−2.31	2.99	0.78	0.58
	Model IV 2	−6.71	5.54	0.53	0.45
	Model IV 3	−1.61	0.96	1.90	0.03
	Model IV 4	−3.15	2.55	1.13	0.07
	PL IV	−0.36	0.28	9.18	0.13
	2S PL IV	−0.98	0.38	3.76	0.07
Italy	Model IV 1	−1.33	1.18	0.50	0.88
	Model IV 2	−1.12	0.77	1.50	0.90
	Model IV 3	−0.83	0.37	2.26	0.92
	Model IV 4	−0.83	0.39	1.95	0.79
	PL IV	−1.25	0.40	3.75	0.56
	2S PL IV	−0.89	0.28	4.26	0.12
Latvia	Model IV 1	3.24	2.06	0.90	0.99
	Model IV 2	0.98	1.05	2.19	0.97
	Model IV 3	0.99	0.47	3.08	0.53
	Model IV 4	0.56	0.47	6.28	1.00
	PL IV	0.70	0.47	3.51	0.52
	2S PL IV	-	-	-	-
Lithuania	Model IV 1	−3.22	2.03	1.38	1.00
	Model IV 2	1.29	2.47	1.10	0.98
	Model IV 3	−0.97	0.62	1.31	0.87
	Model IV 4	−0.09	0.39	3.37	0.36

Table A.1 Continued.

Country	Model	Price Elasticity	Std. Error of Price Elasticity	First-stage F-statistics	Sargan Test (p-Value)
Lithuania	PL IV	−0.08	0.12	1.29	0.18
	2S PL IV	−0.65	0.17	3.99	0.02
Luxembourg	Model IV 1	−1.60	0.70	3.94	0.82
	Model IV 2	−2.96	1.49	2.40	0.99
	Model IV 3	−0.73	0.35	7.67	0.70
	Model IV 4	−1.18	0.42	8.95	0.57
	PL IV	−0.90	0.33	6.41	0.08
	2S PL IV	0.13	0.18	3.28	0.79
Netherlands	Model IV 1	−0.65	1.05	0.20	0.81
	Model IV 2	−0.40	1.62	0.55	0.85
	Model IV 3	0.15	0.42	0.40	1.00
	Model IV 4	0.28	0.66	0.21	1.00
	PL IV	−0.73	0.27	1.37	0.60
	2S PL IV	−0.53	0.06	3.42	0.11
Poland	Model IV 1	−0.39	0.40	8.89	0.10
	Model IV 2	−1.10	0.54	3.22	0.85
	Model IV 3	−0.36	0.13	13.22	0.15
	Model IV 4	−0.55	0.14	11.58	0.30
	PL IV	-0.07	0.06	1.69	0.04
	2S PL IV	0.06	0.17	2.83	0.56
Portugal	Model IV 1	−0.51	1.10	0.13	0.71
	Model IV 2	0.86	0.88	0.37	0.25
	Model IV 3	−0.08	0.47	0.54	1.00
	Model IV 4	0.05	0.39	0.24	0.99
	PL IV	2.73	4.99	2.11	0.17
	2S PL IV	−0.45	0.27	11.37	0.34
Romania	Model IV 1	−2.01	1.85	0.59	0.98
	Model IV 2	−4.58	2.84	1.56	0.97
	Model IV 3	−0.84	0.59	1.75	1.00
	Model IV 4	−0.85	0.89	1.71	0.84
	PL IV	−0.83	0.18	3.39	0.19
	2S PL IV	0.61	0.44	2.60	0.72
Slovakia	Model IV 1	−2.82	4.91	2.25	0.76
	Model IV 2	−10.99	4.99	3.42	0.71
	Model IV 3	−3.20	2.09	3.33	0.06
	Model IV 4	−7.04	1.98	6.76	0.01
	PL IV	0.19	0.20	4.36	0.03
	2S PL IV	−0.68	0.25	10.28	0.12
Slovenia	Model IV 1	−1.33	0.52	5.12	0.97
	Model IV 2	−1.05	0.49	6.51	0.98
	Model IV 3	−1.44	0.25	25.07	0.97
	Model IV 4	−1.25	0.21	23.43	0.53
	PL IV	−1.16	0.27	3.66	0.06
	2S PL IV	−2.23	1.02	2.53	0.21
Spain	Model IV 1	−0.84	1.30	0.81	0.92
	Model IV 2	−0.70	1.18	1.03	0.80
	Model IV 3	−0.36	0.75	2.52	0.85
	Model IV 4	0.04	0.66	0.25	0.65
	PL IV	−0.23	0.49	4.07	0.52
	2S PL IV	−0.32	0.31	3.40	0.60

Table A.1 Continued.

Country	Model	Price Elasticity	Std. Error of Price Elasticity	First-stage F-statistics	Sargan Test (p-Value)
Sweden	Model IV 1	−0.29	12.99	0.13	0.94
	Model IV 2	−1.58	3.50	0.38	0.93
	Model IV 3	0.21	1.38	0.13	1.00
	Model IV 4	−0.26	1.27	0.18	1.00
	PL IV	−0.33	0.12	3.04	0.10
	2S PL IV	−1.12	0.12	1.56	0.08

Table A.2 provides selected set of instrumental and control variables by two-stage post-LASSO IV model for each 23 EU country. The set of possible instrumental variables for a EU country includes other EU countries' HDD and CDD values. In the table, HDD and CDD values of the country XX, which is the country's abbreviation provided in Table A.3 represented as HDD.XX and CDD.XX. For example, the abbreviation of Austria is "AT" and HDD and CDD values of Austrai are represented as "HDD.AT" and "CDD.AT".

Table A.2 Selected set of instrumental an control variables for 23 EU countries by two-stage post-LASSO IV

Country	Instrumental Variables	Control Variables
Austria	HDD.SK, HDD.CZ, HDD.SI, HDD.BG, CDD.BG, CDD.LT, HDD.SE, CDD.FI, HDD.ES, CDD.CY	HDD, UR
Belgium	HDD.NL, HDD.LU, CDD.LU, HDD.FR, HDD.DE, CDD.DE, HDD.DK, HDD.IE, HDD.IE, CDD.AT, CDD.SI, HDD.SK, HDD.HR, HDD.PL, HDD.SE, HDD.ES, CDD.LT, CDD.FI, HDD.BG, CDD.BG, CDD.PT, HDD.MT, CDD.MT, HDD.MT	HDD
Bulgaria	HDD.RO, CDD.GR, CDD.EL, HDD.HU, HDD.HR, HDD.SK, HDD.IT, CDD.IT, HDD.CZ, HDD.MT, CDD.MT, HDD.MT, HDD.PL, HDD.CY, CDD.CY, CDD.LT, HDD.ES, HDD.LU, CDD.LU, HDD.LV, HDD.DK, CDD.DK, HDD.BE, CDD.NL, HDD.FR, HDD.FI, CDD.FI	HDD, GDP, UR
Croatia	HDD.SK, CDD.BG, CDD.DE, CDD.LU, CDD.GR, CDD.EL, HDD.MT, CDD.MT, HDD.MT, CDD.MT, CDD.DK, HDD.LT, HDD.SE, HDD.ES, CDD.ES, HDD.FI, HDD.IE, HDD.IE, HDD.CY, CDD.CY	
Czech Republic	HDD.AT, HDD.DE, HDD.SK, HDD.HU, CDD.HU, HDD.SI, HDD.HR, HDD.LU, CDD.LU, HDD.NL, CDD.NL, HDD.FR, HDD.IT, CDD.IT, HDD.RO, HDD.EE, CDD.FI, HDD.IE, HDD.GR, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.PT, CDD.PT	HDD, CDD, UR
Denmark	CDD.SE, CDD.FI, HDD.SI, CDD.BG, CDD.ES	

Table A.2 Continued.

Country	Instrumental Variables	Control Variables
Estonia	HDD.FI, CDD.FI, CDD.LV, HDD.SE, CDD.SE, CDD.LT, HDD.PL, HDD.DK, CDD.DK, HDD.DE, CDD.CZ, HDD.SK, HDD.AT, CDD.AT, HDD.NL, CDD.NL, CDD.BE, HDD.LU, CDD.LU, HDD.HR, CDD.HR, HDD.SI, HDD.RO, HDD.FR, HDD.BG, CDD.BG, HDD.IE, HDD.IT, CDD.IT, HDD.GR, CDD.GR, HDD.EL, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.CY, CDD.CY, HDD.ES, HDD.BE, CDD.RO	HDD, CDD, GDP, UR
France	HDD.BE, CDD.DE, CDD.AT, HDD.SK, CDD.HU, HDD.PL, CDD.PT, CDD.SE, HDD.MT, HDD.BG, CDD.BG, HDD.FI, CDD.FI	HDD, GDP
Germany	HDD.CZ, HDD.DK, HDD.PL, HDD.SK, CDD.SK, HDD.NL, CDD.NL, HDD.LU, CDD.LU, HDD.BE, CDD.BE, HDD.SI, HDD.HR, CDD.HR, HDD.SE, HDD.LT, CDD.LT, HDD.FR, HDD.EE, HDD.FI, CDD.FI, HDD.IT, CDD.IT, HDD.RO, CDD.RO, HDD.IE, CDD.BG, HDD.GR, CDD.GR, HDD.EL, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.ES, CDD.ES, HDD.PT, CDD.PT	
Greece	HDD.EL, CDD.EL, CDD.BG, HDD.RO, CDD.RO, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.CY, CDD.CY, HDD.IT, HDD.SK, HDD.PL, HDD.DE, CDD.DE, CDD.LU, HDD.BE, HDD.FR, CDD.FR, HDD.NL, CDD.NL, HDD.ES, CDD.ES, HDD.SE, HDD.FI, CDD.FI, HDD.PT, CDD.PT, HDD.IE	
Hungary	HDD.SK, CDD.SK, HDD.HR, HDD.SI, CDD.SI, HDD.CZ, CDD.CZ, HDD.PL, HDD.BG, CDD.BG, HDD.RO, HDD.DE, HDD.IT, CDD.IT, HDD.LT, HDD.LU, CDD.LU, HDD.GR, HDD.EL, HDD.BE, CDD.NL, HDD.FR, CDD.FR, HDD.SE, CDD.SE, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.EE, HDD.FI, HDD.CY, CDD.CY, HDD.IE, HDD.ES, CDD.ES	GDP, UR
Italy	HDD.SI, HDD.MT, HDD.MT, HDD.SK, HDD.BG, CDD.BG, CDD.BE, CDD.PL, CDD.PT, HDD.IE, HDD.IE, HDD.CY	HDD, CDD, GDP, UR
Latvia	HDD.LT, CDD.LT, HDD.EE, HDD.FI, CDD.FI, HDD.SE, CDD.SE, HDD.DK, CDD.DE, HDD.CZ, HDD.SK, HDD.AT, HDD.HU, HDD.NL, CDD.NL, HDD.HR, CDD.SI, HDD.RO, HDD.LU, HDD.BE, CDD.BG, HDD.FR, CDD.FR, HDD.IT, HDD.IE, HDD.IE, HDD.GR, CDD.GR, HDD.EL, CDD.EL, HDD.MT, CDD.MT, HDD.MT, HDD.CY, CDD.CY, HDD.ES, CDD.ES	Price
Lithuania	HDD.PL, CDD.PL, HDD.EE, HDD.FI, HDD.SE, CDD.SE, HDD.DK, CDD.DK, CDD.DE, HDD.SK, HDD.AT	
Luxembourg	HDD.BE, CDD.BE, HDD.NL, HDD.AT, CDD.DK, HDD.SK, HDD.HR, HDD.IE, HDD.IE, HDD.IT, CDD.IT	HDD, GDP

Table A.2 Continued.

Country	Instrumental Variables	Control Variables
Netherland	HDD.BE, HDD.LU, CDD.LU, HDD.FR, CDD.FR, HDD.DE, CDD.DE, HDD.DK, CDD.DK, HDD.CZ, HDD.IE, HDD.IE, HDD.SK, CDD.SK, HDD.SI, CDD.SI, HDD.HR, CDD.HR, HDD.PL, HDD.HU, HDD.IT, CDD.IT, HDD.LV, CDD.LV, HDD.EE, HDD.ES, CDD.ES, HDD.FI, CDD.FI, HDD.BG, CDD.BG, HDD.RO, CDD.RO, HDD.PT, CDD.PT, HDD.MT, CDD.MT, HDD.MT, HDD.GR, CDD.GR, HDD.EL	HDD, CDD, GDP, UR
Poland	HDD.LT, CDD.LT, CDD.CZ, CDD.DE, HDD.SK, HDD.HU, CDD.HU, HDD.AT, CDD.AT, HDD.DK, HDD.HR, CDD.HR, HDD.SE, HDD.EE, CDD.EE, HDD.SI, HDD.FI, CDD.FI, HDD.RO, CDD.RO, HDD.BG, CDD.BG, HDD.LU, CDD.LU, HDD.NL, CDD.NL, HDD.IT, HDD.FR, HDD.GR, CDD.GR, HDD.EL, HDD.IE, HDD.IE, HDD.MT, CDD.MT, HDD.MT, HDD.CY, CDD.CY, CDD.ES	HDD, CDD, GDP, UR
Portugal	HDD.ES, CDD.ES, HDD.FR, HDD.IE, HDD.LU, CDD.LU, HDD.BE, CDD.BE, CDD.IT, CDD.NL, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.HR, CDD.HR, HDD.CZ, HDD.DE, CDD.DE, HDD.SK, CDD.SK, HDD.HU, HDD.DK, CDD.DK, HDD.BG, CDD.BG, HDD.PL, CDD.PL, HDD.GR, HDD.RO, HDD.SE, CDD.SE, HDD.LT, CDD.LT, HDD.EE, HDD.FI, CDD.FI	
Romania	HDD.BG, CDD.BG, HDD.GR, HDD.EL, HDD.SK, CDD.SK, HDD.HR, CDD.AT, HDD.SI, CDD.SI, CDD.CZ, HDD.IT, CDD.IT, HDD.CY, CDD.CY, HDD.MT, CDD.MT, HDD.MT, HDD.LV, CDD.LV, HDD.DK, HDD.LU, CDD.LU, HDD.EE, HDD.SE, CDD.SE, HDD.FI, CDD.FI, HDD.NL, HDD.FR, CDD.FR, HDD.ES, HDD.IE	
Slovakia	HDD.HU, CDD.HR, HDD.CZ, HDD.PL, HDD.DE, CDD.DE, HDD.BG, HDD.IT, HDD.LU, HDD.DK, CDD.DK, HDD.LT, CDD.LT, HDD.BE, HDD.NL, HDD.FR, HDD.SE, CDD.SE, HDD.EE, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.FI, CDD.FI, HDD.IE, HDD.ES, CDD.ES, CDD.CY	HDD, CDD, GDP, UR
Slovenia	HDD.SK, CDD.CZ, CDD.DK, HDD.MT, CDD.MT, HDD.MT, CDD.MT, CDD.GR, CDD.EL, HDD.SE, HDD.ES, CDD.ES, HDD.IE, CDD.CY	HDD, CDD, UR
Spain	HDD.PT, CDD.PT, HDD.FR, CDD.LU, HDD.BE, CDD.IT, HDD.IE, HDD.NL, HDD.MT, HDD.HR, CDD.HR, HDD.CZ, HDD.SK, CDD.DE, HDD.HU, HDD.BG, CDD.BG, HDD.PL, CDD.SE, HDD.EE, HDD.FI, CDD.FI	HDD, GDP
Sweden	HDD.FI, HDD.DK, CDD.DK, HDD.PL, HDD.NL, HDD.SK, CDD.LU, CDD.HR, HDD.IE, HDD.RO, CDD.GR, CDD.EL, HDD.ES, CDD.ES, HDD.MT, CDD.MT, HDD.MT, CDD.MT, HDD.CY, CDD.CY	

Table A.3 Abbreviation of EU countries

Country	Abbreviation	Country	Abbreviation
Austria	AT	Italy	IT
Belgium	BE	Latvia	LV
Bulgaria	BG	Lithuania	LT
Croatia	HR	Luxembourg	LU
Czechia	CZ	Netherlands	NL
Denmark	DK	Poland	PL
Estonia	EE	Portugal	PT
Finland	FI	Romania	RO
France	FR	Slovakia	SK
Germany	DE	Slovenia	SI
Greece	EL	Spain	ES
Hungary	HU	Sweden	SE
Ireland	IE	United Kingdom	UK

APPENDIX B

SUPPLEMENTARY INFORMATION FOR CHAPTER 3

Table Table B.1 and Table Table B.2 display time line of new investment decisions. Specifically, Table Table B.1 shows time of decisions for four LNG regasification terminal projects and Table Table B.2 represents those for 17 pipeline investments.

Table B.1 Timeline of LNG regasification terminal investments

Country	Index	Year				
		1	2	3	4	5
Ireland	1l	Invested	0	0	0	0
Hungary	2l	Invested	0	0	0	0
Greece	3l	Invested	0	0	0	0
Estonia	4l	0	Invested	0	0	0

Table B.2 Timeline of pipeline investments

From	To	Index	Year				
			1	2	3	4	5
Spain	Portugal	1p	0	Invested	0	0	0
Portugal	Spain	1p	0	Invested	0	0	0
Spain	France	2p	Invested	0	0	0	0
France	Spain	2p	Invested	0	0	0	0
Algeria	Italy	3p	0	Invested	0	0	0
Poland	Czechia	4p	0	Invested	0	0	0
Czechia	Poland	4p	0	Invested	0	0	0
Poland	Slovakia	5p	Invested	0	0	0	0
Slovakia	Poland	5p	Invested	0	0	0	0
Austria	Czechia	6p	Invested	0	0	0	0
Czechia	Austria	6p	Invested	0	0	0	0
Greece	Bulgaria	7p	0	Invested	0	0	0
Hungary	Slovenia	9p	0	Invested	0	0	0
Slovenia	Hungary	9p	0	Invested	0	0	0
Romania	Bulgaria	10p	0	0	0	0	0
Romania	Hungary	11p	Invested	0	0	0	0
Hungary	Romania	11p	Invested	0	0	0	0
Hungary	Austria	11p	Invested	0	0	0	0
Estonia	Latvia	13p	0	0	0	0	0
Latvia	Estonia	13p	0	0	0	0	0
Lithuania	Latvia	14p	0	0	0	0	0
Latvia	Lithuania	14p	0	0	0	0	0
Estonia	Finland	15p	0	0	0	0	0
Finland	Estonia	15p	0	0	0	0	0
Poland	Denmark	16p	Invested	0	0	0	0
Denmark	Poland	16p	Invested	0	0	0	0
Poland	Lithuania	17p	0	Invested	0	0	0
Lithuania	Poland	17p	0	Invested	0	0	0

APPENDIX C

SUPPLEMENTARY INFORMATION FOR CHAPTER 4

Table C.1 provides the definitions of abbreviations that are used in this paper.

Table C.1 Table of abbreviations

Abbreviation	Definition
CAPEX	Capital Expenditure
CLLM	Closed-loop Lithium Model
EAS	East Asia
ECA	Europe and Central Asia
EOL	End-of-life
EPA	The United States of America Environmental Protection Agency
ESS	Energy Storage System
EV	Electrical Vehicle
GWP	Global Warming Potential
HHI	Herfindahl-Hirschman Index
ICEV	Internal Combustion Engine Vehicle
IEA	International Energy Agency
IWG	The Interagency Working Group on the Social Cost of Greenhouse Gases
LAC	Latin America and the Caribbean
LCE	Lithium Carbonate Equivalent
LIB	Lithium-ion Battery
M	Million
MEN	Middle East and North Africa
NAM	North America
OPEX	Operational Expenditure
PAC	Pacific
RMP	Raw Material Producer
SAS	South Asia
SCC	Social Cost of Carbon
SSA	Sub-Saharan Africa
US	The United States of America

Equation C.1 represents the non-linear recycling production cost function [75]. w_t and l_t represent input and output of recycling facilities, while $\frac{w_t}{l_t}$ represents efficiency of recycling facilities, respectively. Recall that c^{OPS} and α_t^{TECH} represent initial recycling cost and technological reduction factor.

$$\mathcal{C}(w_t, l_t) = (c^{OPS} - (1 - ln(\frac{w_t}{l_t})^{\rho}))\alpha_t^{TECH} \cdot w_t \tag{C.1}$$

In order to linearize this equation, we propose a set of efficiencies which consists of 0.55% to 0.95% with 0.05% increments. The cost factors in equation C.2 are calculated based on these efficiencies.

$$c^{\text{EFC}} = ln(1 - (\frac{w_t}{l_t})^{\rho}) \tag{C.2}$$