

1-2016

Electric power and the global economy: Advances in database construction and sector representation

Jeffrey C. Peters
Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_dissertations



Part of the [Economics Commons](#), and the [Oil, Gas, and Energy Commons](#)

Recommended Citation

Peters, Jeffrey C., "Electric power and the global economy: Advances in database construction and sector representation" (2016). *Open Access Dissertations*. 694.

https://docs.lib.purdue.edu/open_access_dissertations/694

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

**PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By Jeffrey C. Peters

Entitled

ELECTRIC POWER AND THE GLOBAL ECONOMY: ADVANCES IN DATABASE CONSTRUCTION AND SECTOR REPRESENTATION

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Thomas W. Hertel

Chair

Wally Tyner

Paul Preckel

Joseph Pekny

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Thomas W. Hertel

Approved by: Ken Foster

Head of the Departmental Graduate Program

1/27/2016

Date

ELECTRIC POWER AND THE GLOBAL ECONOMY: ADVANCES IN
DATABASE CONSTRUCTION AND SECTOR REPRESENTATION

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Jeffrey C. Peters

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2016

Purdue University

West Lafayette, Indiana

I dedicate this dissertation and the research it comprises to my loving wife, Shaili,
and my parents, Lynn and John.

...and, of course, Scooter.

ACKNOWLEDGMENTS

I first must acknowledge the wide path to success created by my parents, John and Lynn. They instilled a passion for education and principle from a young age by leading by example – an example I am still learning from today. Nothing I write would do justice to their love and support.

I would also like to acknowledge my wife, Shaili, who has shared life with me throughout this process. She has constantly supported, motivated, and inspired me in my career and in our life, and I am excited to share many more adventures as our future unfolds.

I am thankful for Srinivas Peeta and Dan DeLaurentis who introduced me to academic research when I arrived in the College of Engineering at Purdue University.

In Tom Hertel I found a rare breed of mentor who is a foremost expert and who loves to share his knowledge in a collaborative and supportive way. He always treated me as a colleague and friend, and he continues to elevate my work and life.

Wally Tyner was the first friendly face I found in the department, and he is a testament to timely, interdisciplinary, and policy-relevant research.

The confidence and trust that Joseph Pekny and Angela Phillips Diaz placed on me and in my abilities provided me with unique opportunities to develop into a well-rounded researcher and individual via TEDxPurdueU and GPRI.

Finally, I would like to acknowledge the friendship and inspiration of Yu-Ting Hsu, Freddy Solis, Joseph Louis, and Juan Giraldo who helped make our little town in Indiana a comfortable place to live while providing inspiration to leave our comfort zones as well.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	x
ABBREVIATIONS	xv
ABSTRACT	xvii
CHAPTER 1. INTRODUCTION	1
1.1 Electric Power and the Global Economy	2
1.2 Review of Relevant Literature and Models	5
1.2.1 “Bottom-Up” Versus “Top-Down” Models	6
1.2.2 Input–Output	9
1.2.3 Computable General Equilibrium (CGE)	10
1.2.4 Integrating Top-Down and Bottom-Up	12
1.3 A Path Forward: Adding Electricity-Detail in Computational Equilibrium	13
CHAPTER 2. MATRIX BALANCING WITH UNKNOWN TOTAL COSTS: A NOVEL METHOD FOR GTAP-POWER	17
2.1 Engineering-level Data	19
2.2 Matrix Balancing with Unknown Total Costs	22
2.3 The Structure of the Disaggregation Problem	25
2.3.1 The Disaggregated I–O table	25
2.3.2 The Nature of Economic Data for an Electricity Disaggregation	27
2.4 The Importance of Preserving Economic Relationships in Modeling	28
2.5 Possible Approaches for Unknown Total Costs	31
2.5.1 Kuroda’s Method without a Total Cost Constraint (Kuroda-NC)	32
2.5.2 RAS and Share Preserving Cross-Entropy (SPCE)	33
2.6 An Example in a Simple Electricity Sector Disaggregation	36
2.6.1 Economic Data for the Simple Disaggregation	38
2.6.2 Performance Indicators	39
2.6.3 Comparison of Methods	41
2.7 Conclusions	45
CHAPTER 3. THE RELATIONSHIP BETWEEN MATRIX BALANCING AND MODELING	47
3.1 Some Matrix Balancing Methods	48

	Page
3.1.1 Pro Rata Distribution	49
3.1.2 Minimum Sum of Column Cross-Entropy (MSCCE)	50
3.1.3 RAS	50
3.1.4 Share Preserving Cross-Entropy	51
3.2 Comparison of Construction Methods	51
3.3 Disaggregated Matrices and Numerical Comparison	53
3.4 Economic Implications of Alternative Database Construction Methods	59
3.4.1 Simulation to Highlight the Role of Cost Structure in Modeling	60
3.4.2 Simulation to Highlight the Role of Row Shares	66
3.5 Discussion	68
3.5.1 Disparate Bottom-Up and Top-Down Data	68
3.5.2 Preserving Cost Structure and Row Share	69
3.5.3 Selecting an Appropriate Matrix Balancing Method	69
3.6 Conclusions and Broader impacts	72
CHAPTER 4. CONSTRUCTING THE GTAP-POWER DATABASE	74
4.1 Data	74
4.2 Stage 1: Base and Peak Load Split	77
4.3 Stage 2a: Targeting Levelized Cost Relationships	79
4.4 Stage 2b: Targeting Levelized Cost Sub-Matrices	81
4.4.1 Operating and Maintenance Costs	82
4.4.2 Fuel Costs	82
4.4.3 Capital Costs	83
4.4.4 Own-Use Costs	84
4.4.5 Effective Production Tax	85
4.4.6 Demand and Trade Disaggregation	85
4.5 Results	86
4.5.1 Pro Rata Method for Large Deviations	86
4.5.2 Deviation from Target Levelized Cost Data	89
4.5.3 Main Result	94
4.6 A Look Back at the GTAP Database Construction	101
4.7 Conclusions and Path Forward	103
CHAPTER 5. A DETAILED REPRESENTATION OF THE ELECTRICITY SECTOR FOR ECONOMIC EQUILIBRIUM MODELS	106
5.1 Conceptual Framework	111
5.2 Electricity Production	111
5.3 Capacity Factor Utilization	112
5.3.1 Capacity Factor Flexibility	113
5.3.2 Substitution of Existing Capacity	117
5.3.2.1 Standard CES Substitution	118
5.3.2.2 Quantity-Preserving CES (QCES) Substitution	119
5.3.2.3 Calibrating the Quantity-Preserving CES Parameters	122

	Page
5.3.3 Capacity Factor Utilization Validation	125
5.3.3.1 Baseline Factor Utilization Validation	125
5.3.3.2 Policy-Adjusted Factor Utilization Validation	129
5.3.4 Summary of Capacity Utilization	132
5.4 Capacity Expansion	132
5.5 Joint Capacity Utilization and Expansion Validation	135
5.5.1 Targeted Total Capacity Validation	137
5.5.2 Planning Year Prices Validation	140
5.5.3 Projected Capacity Needs Validation	142
5.5.4 Overall Validation	144
5.6 Summary	146
CHAPTER 6. ON THE ELECTRICITY SECTOR RESPONSE TO THE CLEAN POWER PLAN: CARBON TAX VERSUS INVESTMENT SUBSIDY	148
6.1 Baseline for 2030	151
6.2 Policy Scenarios: Carbon Tax and W+S Investment Subsidy	158
6.3 Policy Scenario Discussion	160
6.4 Conclusions and Future Work	164
CHAPTER 7. CONCLUSIONS AND FUTURE WORK	167
APPENDICES	
Appendix A. Capacity Flexibility, Utilization, and Expansion Analytics: Technology Price Shock	183
Appendix B. Capacity Flexibility, Utilization, and Expansion Analytics: Electricity Demand Shock	188
VITA	192

LIST OF TABLES

Table	Page
2.1 Total electricity production by technology in the United States in 2011, \mathbf{Q}^g (in TWh) and shares. Results may not sum due to rounding. Source: IEA (2014, 2015); Peters (2015).	20
2.2 Levelized cost for electricity production by technology in the United States, l_{it}^0 (2011 US\$/MWh). Source: IEA (2014, 2015); Peters (2015).	20
2.3 Target matrix, \mathbf{A} , for the United States (in millions of 2011 US\$) derived from levelized cost and production data compared to the GTAP totals (\mathbf{U}).	21
2.4 Average rankings of performance indicators for each method. A ranking of 1 is best. Italics show better performing method based on objective type (entropy and SSE-type). Bold shows overall best performing. No superscript, superscript r , and superscript c indicates the metric for deviation from the individual elements, cost structure, and row shares, respectively. The superscript s is the deviation from the uncertain total cost constraint that must be imposed in RAS and Kuroda. Therefore, the metric is only used for SPCE and Kuroda-NC.	42
3.1 Mathematical considerations for comparing matrix balancing methods .	52
3.2 Disaggregated electricity sector for the United States using different approaches (\mathbf{A} , pro rata, MSCCE, RAS, and SPCE) in 2007 US\$. . .	55
3.3 Percentage deviation (mean absolute percentage deviation across inputs and technologies) between the economic relationships before, \mathbf{A} , and after balancing for the United States. Ordering in parentheses.	57
3.4 Capital intensity in the cost structure of technologies after matrix balancing procedures for the United States ($s_{k,t}^V$)	61
3.5 Targeted technology policy: a -30% shock to the price of capital for gas power in the United States	63
3.6 Capital employment across technologies (i.e. row share) after matrix balancing procedures for the United States	66
3.7 Shared input policy simulation: 10% shock to the price of capital in the United States electricity sector	67

Table	Page
4.1 Percent deviation from non-fixed target levelized cost for each generating technology for the United States. The average absolute deviation for non-fixed values is 21.3%.	90
4.2 Percent deviation from non-fixed target shares of levelized cost in the total cost (i.e. cost structure) of each specific generating technology in the United States. The average absolute deviation for non-fixed values is 16.7%.	90
4.3 Percent deviation from non-fixed target relative cost intensity (i.e. row share) normalized by GWh for each levelized cost and generating technology in the United States. The average absolute deviation for non-fixed values is 12.3%.	91
4.4 Example result in the GTAP-Power database: an electricity-detailed disaggregation of the GTAP electricity sector for the United States in 2011. A 24 sector consolidation of the full 63 sectors. Source: Peters (2015)	98
4.5 Deviation between total aggregate inputs to the electricity sector implied by the disaggregate data (used as targets) and the original values in the GTAP ‘ely’ sector (used as consistency constraints) for the United States.	102
5.1 Intensive and extensive margins for electric power at technology and sector-level	108
5.2 Shocks to key drivers of capacity factor from 2002–2012. 2007 is the reference year for GTAPv8. In percentage change from reference year. All are exogenous shifts except \hat{a}^g , which is an output from the model.	126
5.3 Additional shifts to key drivers of capacity factor from 2002 - 2012. 2007 is the reference year for GTAPv8. In percent change from reference year. All are exogenous shifts except \hat{a}^g , which is an output from the model.	130
5.4 Comparison of correlation between observations and model predictions for the baseline and the policy-adjusted validation.	132
5.5 Shocks to key drivers of “targeted total capacity” validation from 2007 to 2018. Total generation, \hat{Q}^g , are exogenously given to target observed total capacity expansion. Policy-adjusted shifts and parameters from Table 5.3 are included in the validation but not shown in this table.	137
5.6 Comparison of correlation between the targeted total capacity (TTC), planning year prices (PYP), and predicted capacity needs (PCN) validations. Each validation has limitations arising from the assumptions, but each lends support to the overall validity of the model. *The low value in correlation here is due to a single outlier.	145

Table	Page
6.1 Connecting the EPA CPP building blocks to model mechanisms determining changes in electricity generation (EPA, 2015).	149
6.2 Shocks to 2030 for baseline scenario based on projections or 2014 observations. Effective taxes from mercury regulation (Table 5.3) and technology (Table 5.5) are also shifted using 2014 levels. Years 2014 (fuel price year) and 2018 (last year of joint validation) used for reference.	152
6.3 Additional shifts to 2030 in addition to baseline shifts in Table 6.2 for carbon tax and W+S investment subsidy policy scenarios. Both policy scenarios choose shocks that meet the total CO ₂ reduction goal of the CPP, 32.0%.	158
6.4 CO ₂ emissions factors for each technology in the United States. EIA data for 2007.	159

LIST OF FIGURES

Figure	Page
1.1 Growth in electricity consumption worldwide. The circle diameter is based on absolute growth, and color intensity is based on percentage growth. Source: EIA (2015)	3
1.2 Global trade in energy products have been growing at a faster rate as compared to non-energy products (in US\$ relative to 1995 baseline). Source: GTAPv8 (Aguiar et al., 2012)	4
1.3 Bottom-up linkages, illustrated by black arrows, show that fixed input costs and exogenously given electricity demand drives the optimal mix of electricity given a detailed representation of electricity supply. On the other hand, top-down (specifically CGE) linkages, illustrated by white arrows, show that supply and demand are in equilibrium (regional and global markets) across all sectors. A new equilibrium (and optimal mix of electricity) is given by some exogenous shock to the initial equilibrium – but input costs are endogenous (not fixed). The dotted lines represent links created by researchers combining top-down and bottom-up methodologies (i.e. hard links and elasticity calibration).	7
2.1 Shares of global electricity generation by technology. About 32% comes from non-fuel-based technologies and would be represented as a portion of total capital in an aggregate electricity sector.	18
2.2 The I–O table for the disaggregated industry (\mathbf{X})	26
2.3 A frequency chart of absolute percentage deviation of the balanced values from the <i>a priori</i> values show that the endogenous weighting scheme of the SPCE formulation is comparable to other exogenous weighting schemes.	36
2.4 A frequency chart of gains from relaxing total cost constraint in CE and SSE-type objectives in terms of difference in $WAP E^r$ and $WAP E^c$	43
2.5 A frequency chart of deviations from the total cost constraint when the constraint is relaxed for CE and SSE-type objectives in terms of difference in $WAP E^s$	44

Figure	Page
3.1 Histograms of percentage deviation between bottom-up and balanced data in each region for both cost structure (a) and row share (b) - where deviation is the absolute percentage deviation averaged across inputs and technologies in each region.	58
3.2 Production structure for the simple PE model of the representative electricity sector	62
3.3 Histogram of absolute percentage deviation from bottom-up model results and balanced data model results from a -40% shock to the price of gas in each of the 129 GTAP regions.	65
3.4 Considerations for selecting an appropriate matrix balancing method – insights from algorithms. These only hold when no additional informational constraints are present.	71
3.5 Considerations for selecting an appropriate matrix balancing method - Insights from modeling. The corresponding insight from the algorithm (Figure 3.4) are in parentheses	71
4.1 Shares of global electricity generation from different technologies in base and peak (green cut-out) load.	78
4.2 Share of total GWh which converge using the SPCE procedure with the bound of deviation from target LCOE data	88
4.3 Histogram of deviation between estimated levelized cost and target levelized costs for all regions. For each distribution, the deviation of the median from one indicates bias and larger standard deviation indicates larger deviation between the disaggregate data and original GTAP data. Note: l_{ct}/l_{ct}^0 plotted on a log-scale.	93
4.4 A histogram comparing deviation between estimated levelized costs and target levelized costs for OECD countries and non-OECD countries. The larger mass of OECD regions around one indicates a closer match between disaggregate data and original GTAP data. Note: l_{ct}/l_{ct}^0 plotted on log-scale. Non-OECD counts (1,408 non-fixed values) are scaled to OECD counts (1,111 non-fixed values) for comparison.	95
4.5 The share of GWh produced by each technology in GTAP-Power for the same ten largest fuel-based electricity sectors plus Brazil as Figure 4.6 (ordered left to right by share of non-fuel based technology share). The non-fuel based technologies represented with diagonal lines (i.e. Nuclear, Wind, HydroBL, Other, HydroP, and Solar) are only implicitly represented by ‘Capital’ in the original GTAP database.	99

Figure	Page
4.6 The share of domestic and imported fuels used in the electricity sector for the ten largest fuel-based electricity sectors (ordered left to right by import intensity) plus Brazil. Both the fuel and source composition differ greatly between countries.	100
4.7 Korea produces more than 32% of its electricity from imported fuels. These charts show the composition of Korea's coal and gas imports. . .	101
5.1 The white part of the matrix represents the partial equilibrium model described in Chapter 5 with supply curves for coal, gas, oil, O&M, and capital. The light gray areas signify the full GTAP CGE linkages which have greater sectoral detail as well as supply and demand schedules across all sectors, households, and government rather than simple supply curves. Integrating the electricity-detail in GTAP requires introducing the full sectoral detail for inputs as well as the demand for electricity for each user (i.e. firms, household, and government)	110
5.2 Total electricity generation is determined by the dual capacity utilization and capacity expansion mechanisms and their interdependency. The two mechanisms are linked by the capacity factor for existing and new capacity (\hat{c}_t), returns to capacity ($\hat{p}_{k,t}$), and the net change in capacity (\hat{q}_t^c). . . .	112
5.3 Annual capacity factors from 2002–2012 EIA (2015b). The slope of the trend lines indicate the technology's flexibility in the face of changing economic conditions over time. Black lines represent the flexible technologies, while the gray trend lines represent the inflexible technologies.	114
5.4 Generic supply curve for a flexible electricity generating technology. An inflexible technology would be a vertical line intercepting the x-axis at the capacity factor value.	115
5.5 Capacity factor supply curves for flexible electricity generating technologies. Data points represent results from shocks to other substitutable technologies to shift demand for the relevant technology. The selection of the shocks were designed to map out the response over a wide range of possible shocks.	116
5.6 Production structure for capacity utilization. Composite sectors are in <i>italics</i>	123
5.7 Fuel prices per MWh of electricity produced (nominal dollars)	127

Figure	Page
5.8 Validation of capacity factor portion of model. Model results are represented by large markers and are not connected because they are each shifted separately from the 2007 base year. Observed values are gray, dashed lines.	127
5.9 Refinement of capacity factor validation using additional insights. Model results are represented by large markers and are not connected because they are each shifted separately from the 2007 base year. Observed values are gray, dashed lines	131
5.10 The “targeted total capacity” validation controls for total capacity expansion. Model results are for each year are changes from the 2007 baseline.	138
5.11 Gas prices unexpectedly fell beginning in 2008, so due to assumption of input prices at service year instead of input price at planning year, the model over-predicts expansion of gas capacity. The model predicts a more immediate turn from new coal capacity while there is some lag in the observed data.	139
5.12 The model does well to predict capacity expansion in renewable power.	141
5.13 Gas capacity expansion in the “planning year price” validation. Using planning year input prices corrects for the over-prediction using service year prices, but does not allow for year-to-year planning adjustments. Model results are for each year are changes from the 2007 baseline. . .	142
5.14 Model results for each year are changes from the 2007 baseline.	143
5.15 Model results for each year are changes from the 2007 baseline.	144
6.1 CO ₂ emissions by source for the US electricity sector from 1989 to 2010, and as total emissions from 2011 to 2013 (EIA, 2015b). We observe decreasing CO ₂ intensity from 2007 onward (A). The total generation projection for 2030 is based on EIA AEO 2015 (EIA, 2015a), and the two total CO ₂ projections are based on CO ₂ intensity projections using linear regressions with different periods of the emission data (1989–2013 and 2007–2013).	151
6.2 Contributions to total CO ₂ emissions by fuel-type in the United States in 2005 and model projections for 2030. Emissions from coal and oil reduce by 48% and 60%, respectively, but is partly offset by a 72% increase in emissions from gas power. CO ₂ emissions for other power technologies is small.	154

Figure	Page
6.3 Capacity utilization in 2018 and 2030 are roughly similar due to similar fuel price shifts (dissimilar O&M price shifts). This results in returns to capacity in 2018 and 2030 are roughly similar except for technologies that have difficulty expanding economically due to policy (i.e. coal and oil) or resource and regulatory constraints (i.e. nuclear and hydro).	155
6.4 The shares of additional nameplate capacity for 2007–2018 and for 2007–2030 are roughly similar. Areas to scale.	156
6.5 A carbon tax reduces capacity utilization for fossil fuels from the baseline based on relative carbon content (1, 2) while a wind and solar investment subsidy impacts utilization for all other flexible technologies in a more uniform manner (1, 2, 3).	161
6.6 Absolute returns for fossil fuels are reduced by the carbon tax and less impacted by the wind and solar investment subsidies (4). Returns to other zero emissions technologies that not subsidized are hurt by the investment subsidy scenario (5).	162
6.7 Power generation in 2030. A carbon tax allows some zero emission technologies (i.e. boxes with vertical lines) to expand slightly while these are reduced in the W+S subsidy case. The carbon tax allows for greater switching from gas to coal (CPP building block 2) as compared to the W+S subsidy case. As expected larger growth in wind and solar is observed with the subsidies.	163

ABBREVIATIONS

AEO	US Energy Information Administration Annual Energy Outlook
CAMR	Clean Air Mercury Rule
CES	constant elasticity of substitution
CGE	computable general equilibrium
CO ₂	carbon dioxide
CPP	Clean Power Plan
EDGE	electricity-detailed general equilibrium
EIA	US Energy Information Administration
EPA	US Environmental Protection Agency
GEMPACK	General Equilibrium Modeling Package
I-O	input-output
IAM	integrated assessment model(ing)
IEA	International Energy Agency
LCOE	levelized cost of electricity
LNG	liquefied natural gas
MATS	Mercury and Air Toxics Standards

MNL	multinomial logit
MSCCE	minimum sum of column cross-entropy
MW	Megawatt(s)
MWh	Megawatt-hour(s)
GHG	greenhouse gas
GTAP	Global Trade Analysis Project
GW	Gigawatt(s)
GWh	Gigawatt-hour(s)
O&M	operating and maintenance
OECD	Organisation for Economic Co-operation and Development
PE	partial equilibrium
QCES	quantity-preserving constant elasticity of substitution
RAS	R and S bi-proportionate adjustment method
SPCE	share preserving cross-entropy
T&D	transmission and distribution
TWh	Terawatt-hour(s)
US	The United States of America

ABSTRACT

Peters, Jeffrey C. PhD, Purdue University, May 2016. Electric Power and the Global Economy: Advances in Database Construction and Sector Representation. Major Professor: Thomas W. Hertel.

The electricity sector plays a crucial role in the global economy. The sector is a major consumer of fossil fuel resources, producer of greenhouse gas emissions, and an important indicator and correlate of economic development. As such, the sector is a primary target for policy-makers seeking to address these issues. The sector is also experiencing rapid technological change in generation (e.g. renewables), primary inputs (e.g. horizontal drilling and hydraulic fracturing), and end-use efficiency. This dissertation seeks to further our understanding of the role of the electricity sector as part of the dynamic global energy-economy, which requires significant research advances in both database construction and modeling techniques. Chapter 2 identifies useful engineering-level data and presents a novel matrix balancing method for integrating these data in global economic databases. Chapter 3 demonstrates the relationship between matrix balancing method and modeling results, and Chapter 4 presents the full construction methodology for GTAP-Power, the foremost, publicly-available global computable general equilibrium database. Chapter 5 presents an electricity-detailed computational equilibrium model that explicitly and endogenously captures capacity utilization, capacity expansion, and their interdependency – important aspects of technological substitution in the electricity sector. The individual, but interrelated, research contributions to database construction and electricity modeling in computational equilibrium are placed in the context of analyzing the US EPA Clean Power Plan (CPP) CO₂ target of 32 percent reduction of CO₂ emissions in the US electricity sector from a 2005 baseline by 2030.

Assuming current fuel prices, the model predicts an almost 28 percent CO₂ reduction without further policy intervention. Next, a carbon tax and investment subsidies for renewable technologies to meet the CPP full targets are imposed and compared (Chapter 6). The carbon tax achieves the target via both utilization and expansion, while the renewable investment subsidies lead to over-expansion and compromises some of the possibilities via utilization. In doing so, this dissertation furthers our understanding of the role of the electricity sector as part of the dynamic global energy-economy.

CHAPTER 1. INTRODUCTION

The electricity sector is crucial to the global economy. Electricity production and consumption have been identified as important indicators and correlates of economic development. Payne (2010) provides a comprehensive survey of research in individual countries related to the causality between electricity production and economic growth. The electricity sector is also a major consumer of fossil fuels. In fact, national electricity sectors are major drivers of energy exports for some resource-intensive countries (especially natural gas and coal exporters). The electricity sector also accounts for roughly 40% of greenhouse gas (GHG) emissions in the United States and 30% worldwide. It is no surprise then, that many countries have targeted the electricity sector with policies meant to achieve both economic growth and GHG mitigation targets. The sector is also experiencing rapid technological change in generation technologies (e.g. advances in renewable energy), energy inputs (e.g. horizontal drilling and hydraulic fracturing of shale deposits), and end-use efficiency. Many important questions in global economic analysis revolve around the electricity sector. How can we answer electricity-related questions when the surrounding technology and economy are evolving at such a rapid pace? The specific question investigated here is: how does the electricity sector respond to a carbon tax versus regulation? A tax is known to be economically efficient, but political intractability has instead led to alternative strategies (e.g. the Clean Power Plan). What does this mean for the electricity sector? This is just one of the many possible applications of the body of work presented here.

1.1 Electric Power and the Global Economy

Electricity production and consumption have been linked as important indicators and correlates of economic development. In a survey of over 100 countries representing 99% of the global economy, Ferguson et al. (2000) conclude that as a country's wealth increases, the proportion of electricity use to total energy use increases as well. They also show that there is a stronger correlation between electricity use and wealth than between total energy use and wealth. Beyond correlation, Payne (2010) provides a comprehensive survey of research in individual countries related to the causality between electricity and economic growth. Apergis and Payne (2011) show that for low-income countries there is a unidirectional correlation from electricity production to economic growth and a bidirectional causality in higher income countries. Because of the evidence supporting this link, electricity production, sources, and access are used as economic and environmental indicators by the World Bank (2012).

Most of the electricity growth internationally has come from developing countries. The International Energy Agency (IEA) estimates that electricity output has increased 81% worldwide from 1990 to 2010, and growth was higher in non-OECD countries compared to OECD countries, 152% and 42% respectively (IEA, 2011) (see Figure 1.1). Rapid electricity growth in developing countries increasingly relies on coal, while natural gas has increased in industrialized countries as a result of fuel switching efforts (Quadrelli and Peterson, 2007). In the United States, fuel switching to gas has been driven to a large extent by falling domestic gas prices. In light of the wide range of electricity generating technologies available, these trends are likely driven by cost and policy considerations (i.e. relatively low costs of these fuels as compared to other types of generation).

While electricity itself is largely produced and consumed domestically, the fuel inputs (e.g. oil, coal, gas) are increasingly traded internationally (See Figure 1.2). Petroleum only accounts for a small share of electricity generating fuels but is the foremost traded commodity globally in US\$ (United Nations, 2012). 'Hard' coal accounts for about 40% of global electricity production and is increasingly being

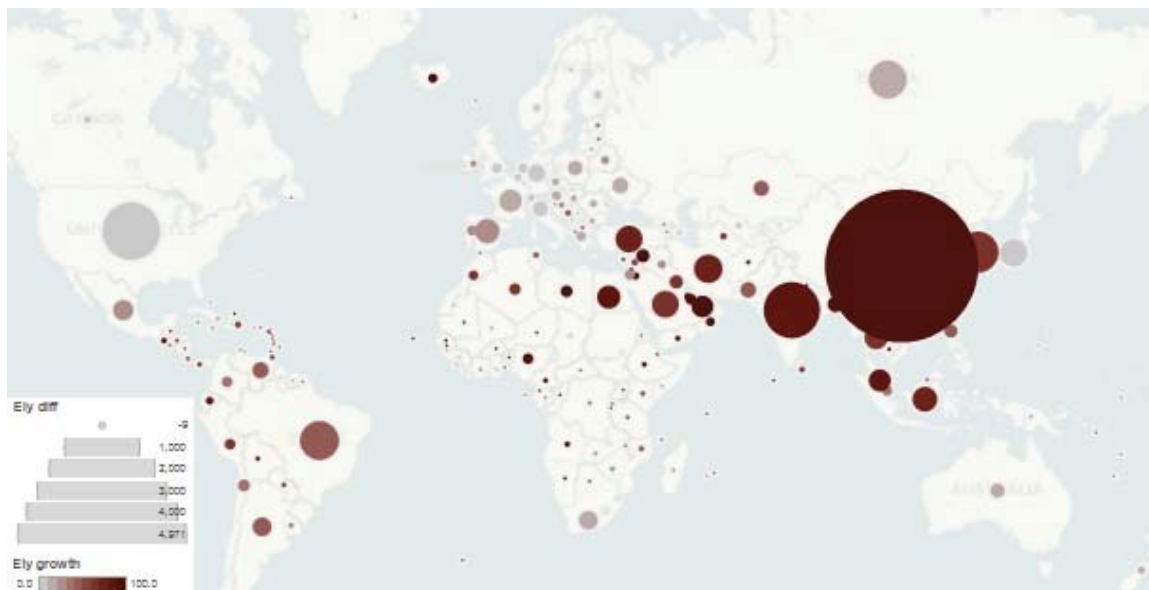


Figure 1.1.: Growth in electricity consumption worldwide. The circle diameter is based on absolute growth, and color intensity is based on percentage growth. Source: EIA (2015)

traded internationally (Haftendorn and Holz, 2010; United Nations, 2012). Also, recent developments in shale oil and gas extraction have instigated serious discussions around the world centered on expanding global trade and the resulting implications for economic growth, GHG policy, and energy security (Egging et al., 2008; Paltsev et al., 2005).

Environmental policies in the electricity sector and technological improvements in energy production may result in changes in global electricity fuel trade. Many developed countries have taken a stance against high carbon emitting (namely coal-fired) generation or for renewable generation technologies. McCollum et al. (2014) show that climate mitigation policies may impact fossil-fuel consumption, trade, and prices over time. For example, despite having the largest share (28%) of global coal reserves (EIA, 2015), the United States recently reduced the ceiling on allowable carbon emissions of new power plants (Obama, 2013; EPA, 2015), effectively ensuring no new coal plants are constructed unless currently uneconomic carbon capture and sequestration measures are implemented. Thus, the inescapable

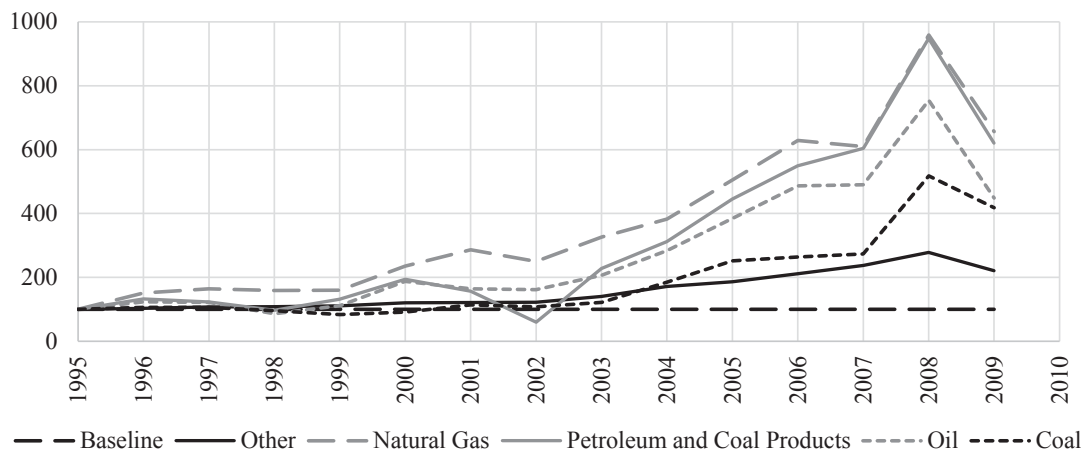


Figure 1.2.: Global trade in energy products have been growing at a faster rate as compared to non-energy products (in US\$ relative to 1995 baseline). Source: GTAPv8 (Aguiar et al., 2012)

question is: where will all the US coal go and what will happen to the stranded assets? In the United States, horizontal drilling and hydraulic fracturing has expanded technically recoverable energy reserves (i.e. oil and gas). This technological shock allowed the United States to surpass Russia as the world's top gas producer and is projected to be the top oil producer by 2020 (IEA, 2011). With US natural gas prices roughly one-third that of world prices, exports appear opportune (Levi, 2012). The inescapable question here is: how will the trade of these fuels shift electricity production in the United States and around the world? These questions are not limited to the United States; many other countries are facing domestic environmental policies and shale oil and gas opportunities (Kuuskraa et al., 2013).

A significant shift in global trade of electricity fuels may have sweeping implications concerning global carbon emissions. This requires investigating two important economic margins of adjustment: bilateral sourcing of traded goods and substitutability amongst energy inputs to electricity generation. First, disparate regional allocations, technologies, and prices have an impact on climate mitigation (Peters and Hertwich, 2008; Chen, 2009; Bushnell and Chen, 2009) when considering cross border trade. It is necessary to analyze domestic production along with sources

of imports and exports to capture such global implications. Second, ignoring several operational considerations, electricity can be considered a homogeneous output with various heterogeneous input technologies. Substituting fuels requires both economic incentives and physical potential, which are dependent on the characteristics of the technologies and operation in the electric power sector (Delarue and D'haeseleer, 2008). In the context of carbon mitigation, emissions from electricity production can be reduced by decreasing output, increasing carbon efficiency of existing electric power technologies, and/or substituting fuels to production (Soytas et al., 2007; Haszeldine, 2009).

This dissertation aims to further our understanding of the electricity sector as part of the global economic landscape. The remainder of this chapter reviews relevant literature and existing methodologies and then proposes specific research advances to meet this aim.

1.2 Review of Relevant Literature and Models

Innumerable studies attempt to shed light on the impact of economic shocks on the electricity sector. The bulk of attention in this particular review is placed on methodologies which have a detailed representation or focus on electricity as well as regional-global linkages. Even this subset of the relevant research entails a very extensive literature. Models which have received recent attention in the context of the research objectives of this study include: “bottom-up”, input–output, and computable general equilibrium (CGE) models. The remainder of this section is devoted to describing the structure, relevant research, advantages, and limitations of each approach. It is easily seen in this review that climate change and carbon mitigation are the primary foci for researchers studying global impacts of electricity trade and policy.

1.2.1 “Bottom-Up” Versus “Top-Down” Models

The distinction between “bottom-up” and “top-down” models is best described by visualizing a hierarchy of the economy with the macro-economy at the top and individual decision makers (e.g. agents, firms) near the bottom (see Figure 1.3). Similarly, Hourcade et al. (2006) describe an “ideal” energy-economic model as having the following three characteristics: i) technological explicitness, ii) macroeconomic completeness, and iii) microeconomic realism (Hourcade et al., 2006, Figure 1). Bottom-up models have the first two characteristics by pursuing detail at the technology level and building up to analyze the macroeconomic impacts. Top-down models have the latter two characteristics by starting with the macro-economy and economically-consistent linkages between aggregate economic agents to analyze impacts on the various sub-sectors (e.g. electric power).

Bottom-up models, also termed engineering (partial equilibrium) models, use a wide array of information, technological parameters, and discrete decision-making to explicitly represent both the demand and supply patterns in a particular sector. Such detail for electricity includes, but is not limited to the following dimensions of electric power supply and demand:

- electric appliance-level (e.g. refrigerator, air conditioning, electric vehicle) consumption specifications,
- time-of-use (e.g. night/day, seasonal) changes in demand load patterns,
- dynamic capacity, technical, and resource availability constraints in supply load patterns,
- base, peak, and intermediate power profiles,
- transmission and distribution costs and power losses, and
- cost and technology paths for existing and new technologies which reflect learning, location, specific regulations/policy, and other information (Neij, 2008; Lanz and Rausch, 2011).

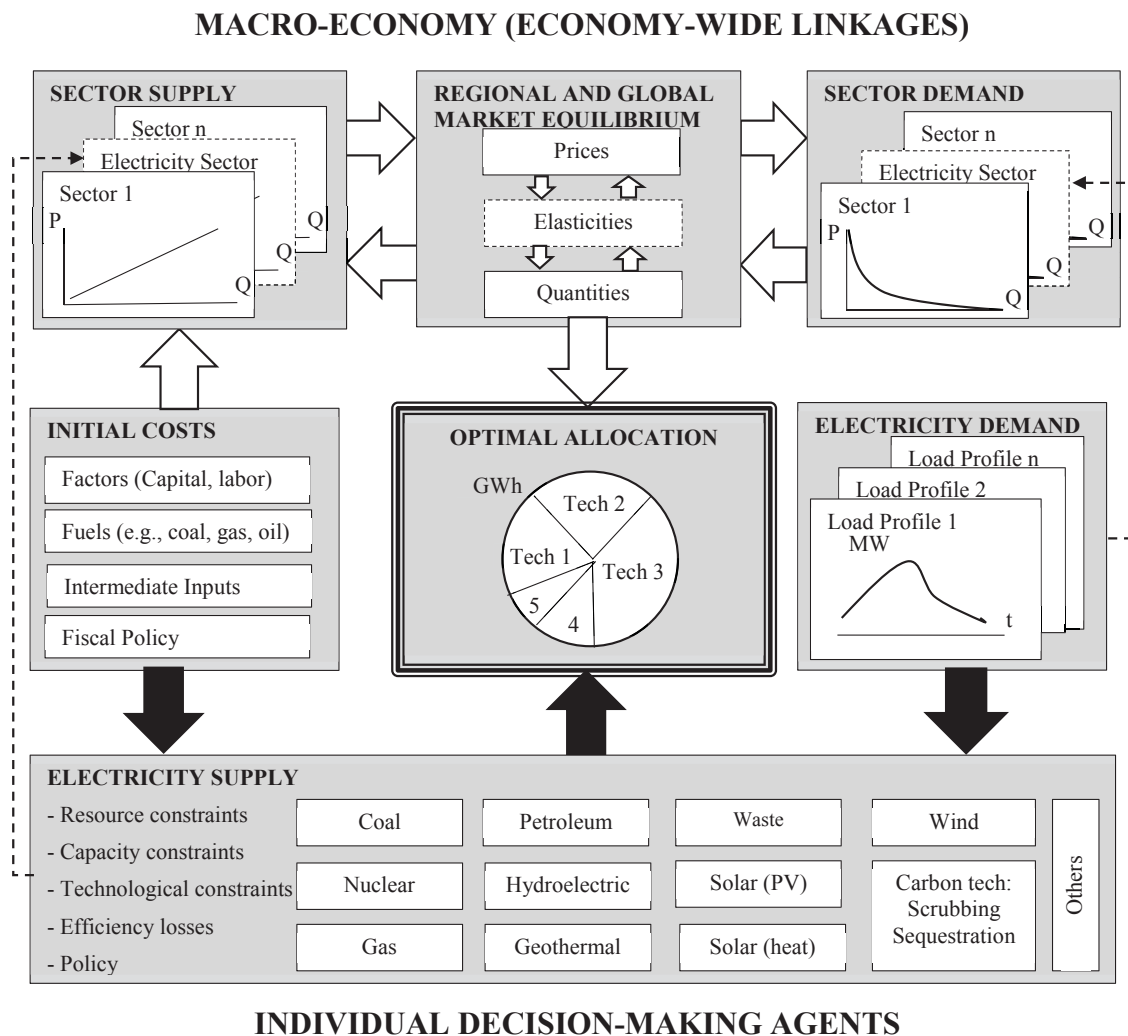


Figure 1.3.: Bottom-up linkages, illustrated by black arrows, show that fixed input costs and exogenously given electricity demand drives the optimal mix of electricity given a detailed representation of electricity supply. On the other hand, top-down (specifically CGE) linkages, illustrated by white arrows, show that supply and demand are in equilibrium (regional and global markets) across all sectors. A new equilibrium (and optimal mix of electricity) is given by some exogenous shock to the initial equilibrium – but input costs are endogenous (not fixed). The dotted lines represent links created by researchers combining top-down and bottom-up methodologies (i.e. hard links and elasticity calibration).

All of the above factors influence the selection of inputs and technologies used for the ultimate production of electricity. The optimal selection is based on some given objective (e.g. cost minimizing, welfare maximizing). This type of activity-based analysis contrasts with the more aggregated production functions utilized in the top-down models. In top-down models, the aggregate implications of cost-minimizing behavior are summarized based on cost shares and elasticities of substitution amongst inputs or aggregate technologies.

There are two basic types of bottom-up models: optimization and simulation. Optimization models attempt to uncover the least cost or maximum surplus path of producing electricity in a partial equilibrium framework (i.e. given exogenous input prices). The MARKAL model offers a representative example of these models (Loulou et al., 2004). Alternatively, simulation models attempt to capture the behavior of individual economic agents (e.g. consumers, firms) involved in production explicitly and may, or may not strictly follow specific economic assumptions (e.g. profit maximization). In any case, they do not aim to maximize aggregate economic surplus in the same way as do the optimization models. Examples include models which aim for behavioral realism (Jaccard et al., 1996) and models which connect disparate engineering-type systems via market mechanisms (Hodge et al., 2011).

Bottom-up models provide a great deal of sector-specific information, which make them a powerful tool in modeling the electric power sector. However, they typically have limited ability to capture economy-wide interactions, notably the lack of endogenous price effects. Further, the models are often not linked to global trade, which has been cited as an important aspect in analyzing economic shocks related directly or indirectly (via energy inputs) to the electric power sector and the impact of electricity production (e.g. CO₂ emissions). For instance, Bataille et al. (2006) highlights the importance of endogenous price feedbacks from market equilibrium which may be ignored in strictly bottom-up models. Also, Chen (2009) and Bushnell and Chen (2009) use a simulation model to show emission leakage across borders in the context of carbon trading schemes. While bottom-up models may capture

some trade pattern-related issues, other sectors which consume the energy inputs and electricity are largely not considered and bilateral sourcing ignored, limiting the ability to determine several economy-wide impacts. As trade in electricity fuels increases, these effects may become increasingly relevant.

Top-down models incorporate regional differences and trade effects in determining the impact of economic shocks in the electricity and energy sectors, sacrificing operational detail for intersectoral, interregional, and economy-wide consistency.

1.2.2 Input–Output

Input–output (I–O) analysis is widely used to analyze inter-sectoral interactions. They typically have limited economic detail (e.g. fixed input–output coefficients, exogenous input prices, unlimited factor supply); however, they can provide a rich description of how different sectors of the global economy are interrelated. Peters and Hertwich (2008) use the Global Trade Analysis Project (GTAP) database for 87 countries and 57 industries to drive a multi-regional input–output table (MRIO). They use this to conclude that emissions involved in trade have significant implications in participation and effectiveness of global climate policies due to different regional trade and emission impacts from specific policy designs. Yunfeng and Laike (2010) explore CO₂ emission leakage in China using I–O analysis and conclude that developed countries transferred a large amount of pollution by offshoring manufacturing and other carbon-intensive industries to China. Although the study does not specifically focus on electricity, they claim that every unit of energy used in China results in more CO₂ emissions than in developed countries because of coal-intensive electricity generation ($\approx 80\%$ of total production). Weisser (2007) investigates electric power specifically using life-cycle assessment, a variant of I–O which attempts to capture GHG emissions over the production supply chain (resource exploration to electric power production to waste management). Significant emissions (e.g. up to 25%) up- and downstream of domestic electric power production process may occur outside of

legislative boundaries, offsetting policy effectiveness. These three studies show the importance and effectiveness of MRIO analysis in regional boundary considerations by capturing the electric power supply chain and offshoring of energy-intensive industry.

I–O does a good job of capturing intersectoral and interregional characteristics of the global economy in terms of quantities. In fact, West (1995) describes I–O models as “bread and butter” models for regional economic impact analysis. However, I–O models also have significant limitations in the case of electricity production and global fuel trade including the use of linear functions and exclusion of price effects (similar to bottom-up models). Various extensions of MRIO models overcome some, but not all, of these limitations (e.g. full price effects). These limitations are significant in the electric power sector, due to the extent of energy use and substitutability of energy inputs.

1.2.3 Computable General Equilibrium (CGE)

Limitations of bottom-up and I–O models point to CGE models to include a detailed representation of electricity with regional-global linkages. CGE models determine both price and quantity changes endogenously in the wake of a given exogenous shock to the global economy. They are built on I–O databases and therefore capture intersectoral and interregional linkages; however, unlike I–O models, supply and demand are equated through a market clearance condition. Production can be characterized in a multitude of ways including nested Cobb-Douglas, CES, or Leontief functions (as opposed to only Leontief-type in I–O) which can add sector-specific details such as substitutability of inputs (West, 1995). CGE also allows for full price effects (e.g. exogenously determined input prices) whereas bottom-up and I–O typically assume prices are exogenous. Hazilla and Kopp (1990) and Bergman (1991) conclude general equilibrium impacts, such as input prices, output prices, and allocation of resources in the economy, can be “significant and pervasive” in the context of environmental policy – a highly relevant finding in light of the theme of this

dissertation. Substitutability of inputs, unique production structures, and strength in capturing global trade patterns highlight CGE as an ideal foundation for many applications related to global impacts focusing on the electricity sector.

However, DeCanio (2003) offers sobering critiques of CGE models used in climate change analysis (although debate remains in regard to the gravity of these concerns (Rutherford, 2005; Koomey, 2005)). First, CGE models require many parameters and data (e.g. elasticities, value shares) which are difficult to estimate beyond “best guesses” of experts. Many researchers question the reliability of CGE models because of their sensitivity to these parameters. Second, CGE models are rarely validated against past experience, leading many researchers to perceive these as merely illustrative and not predictive tools. Beckman et al. (2011) compare historical and GTAP-E (an energy-environmental CGE model) based predictions of petroleum price distributions and conclude that the original, not econometrically estimated, GTAP-E parameters were too elastic. This leads them to incorporate more recently econometric estimations of energy demands, after which the model’s performance is significantly improved. Third, the top-down nature of CGE models precludes incorporating detailed information at the sector level. Considerations of the electric power sector (e.g. technological feasibility, resource availability, dynamic operations), which are explicitly incorporated in bottom-up models, are only implicitly addressed in CGE models (Williams et al., 2012).

For example, the GTAP database is the predominant database for global CGE modeling; however, there is only a single electricity sector which encompasses production, collection, and distribution. Several researchers have independently disaggregated the sector into various generating technologies and incorporated specialized production structures for the electricity sector to mimic bottom-up electric power considerations into the top-down nature of global CGE models (e.g. Burniaux and Truong (2002), Paltsev et al. (2005), Wing (2006), Pant (2007), Château et al. (2014)). Unfortunately, the disaggregation methods remain largely undocumented and sector-specific detail untested against real-world observations. The foremost

challenges in modeling electricity in the CGE framework are: i) data and methods to construct an electricity-detailed CGE database, ii) a way to explicitly and reliably incorporate key sector-specific considerations, and iii) methods for model validation.

1.2.4 Integrating Top-Down and Bottom-Up

As mentioned above, CGE modelers have increased the detail (and complexity) of their production structures for certain commodities to capture unique technological considerations of the sector. To support these efforts, insights from bottom-up models can be used to enhance both the parameters and data. On the parameter side, Schäfer and Jacoby (2005) adjust parameters in a top-down CGE model to reflect behavior predicted by a bottom-up model (MARKAL). Similarly, Kiuila and Rutherford (2013) demonstrate that estimates of elasticities of substitution from historical data may not be valid under potential technological and policy changes. They propose methods for approximating the elasticity of substitution between technologies from step functions characteristic of bottom-up models.

Other researchers have explored methods to explicitly combine bottom-up and top-down models in a single framework. The MARKAL-MACRO model (Manne and Wene, 1992) combines the technological detail of the energy sector in the MARKAL model (Loulou et al., 2004) with a simple general equilibrium from the MACRO model (Manne and Richels, 1992) via exchange of energy output and energy cost variables between the respective models. Similarly, Schäfer and Jacoby (2006) create a single framework by exchanging prices, demand, and modes shares from a CGE model with substitution elasticities from the MARKAL model. Böhringer and Rutherford (2008) describe a complementarity formulation of the energy sector which combines bottom-up technological detail and top-down economic considerations within a single mathematical model. Because of complexity issues which may make the approach intractable, Böhringer et al. (2009) decompose the original complementarity formulation into separate top-down and bottom-up models, which

are solved independently, and uses an iterative process termed “sequential calibration” to obtain convergence of results from the two models.

1.3 A Path Forward: Adding Electricity-Detail in Computational Equilibrium

The primary purpose of this research is to advance our understanding of the role of electricity in the global economic landscape by focusing on: i) economic and technological considerations of the electricity sector, namely the substitutability between specific generation technologies and ii) economy-wide linkages.

While bottom-up models have been enormously successful in the former goal, the lack of price feedbacks and global trade rigor/richness leads to limitations in analyzing intersectoral, interregional, and economy-wide effects. Also, by not capturing full price effects or substitutability of generation technologies, input–output analysis suffers from an inability to capture specific economic and technological considerations. These limiting factors compromise the ability of both models to adequately model the global economic implications of regional shocks considering electricity (e.g. trade shifts, welfare changes, energy security perspectives).

CGE models capture full price effects with the intersectoral and interregional linkages necessary for analyzing the stated goals of the research. Rigor and richness in trade patterns (e.g. bilateral sourcing) provide an ideal platform for understanding global impacts such as GHG emissions. Adding first-order sector-specific detail would allow for greater credibility to CGE models without sacrificing their benefits.

Thus, the aim of the dissertation is to construct a model with a technologically-rich representation of the electricity sector which can comprise part of a global CGE model with fully endogenous prices and production (i.e. feedbacks in the dynamic energy economy). This involves two main steps: i) constructing a technologically-rich electricity sector within a “top-down” economic database and ii) development of a computational equilibrium model which captures important operational characteristics of the electricity database in a manner consistent with

existing global economic analysis. These two broad steps and their individual contributions to literature are discussed in detail in Chapters 2–4 and 5, respectively. In Chapter 6 the computational equilibrium model is used to answer the question: how does the US electricity sector respond to a carbon tax versus regulation? Chapter 7 concludes.

Detailed in Chapters 2–4, the database construction step, is a disaggregation of the electricity sector in the Global Trade Analysis Project (GTAP) database, an applied general equilibrium database of the world economy constructed from national input–output tables, trade, macroeconomic, and trade protection data from several sources. As discussed earlier, the main GTAP database, however, only has one sector which encompasses the production, collection, and distribution of electricity. This aggregate electricity sector is disaggregated into the following new sectors: transmission and distribution (T&D), nuclear, coal, gas, oil, hydroelectric, wind, solar, and ‘other’ power technologies. Gas, oil, and hydroelectric power are further differentiated by load type: base and peak. Chapter 2 introduces the engineering-level data to inform the disaggregation and a novel method that best preserves economic information considering specific aspects of the electricity disaggregation problem. The method is useful for applications outside electric power as well. Chapter 3 discusses the linkages between various matrix balancing methods and modeling results and concludes that preserving certain economic relationships implied by the engineering-level data is of utmost importance, which lends confidence to the methodology. Chapter 4 documents additional components of the complete GTAP-Power disaggregation.

Chapter 5 presents a model that uses the GTAP-Power database as a starting point and characterizes how electricity generating technologies might evolve in response to technological advances and policies. Substitution between electric power generating technologies results from two distinct mechanisms: i) constructing new capital, termed capacity expansion, and/or ii) increasing or decreasing operations of existing capital, termed capacity utilization. Long-term returns on capital investment in

electric power technologies drive expansion, while utilization is the substitution of production from existing capacity, also called fuel-switching, in response to prevailing economic conditions, namely fuel prices. The two mechanisms are interrelated in that capital rents partly depend on how much generation is produced per unit of capacity (i.e. capacity factor) and short-term utilization changes may be counterbalanced by long-term expansion. The model posed here integrates these two mechanisms of electricity generation within a **partial equilibrium** framework conducive to extension into a full computable general equilibrium framework. Importantly, this representation of the electricity sector is validated against observations. A corollary contribution to the economic modeling literature is a novel implementation of a variant of the constant elasticity of substitution (CES) production specification which ensures the quantity of aggregate electricity production (i.e. GWh) is equal to the sum of the input quantities (i.e. GWh from individual technologies) (van der Mensbrugge and Peters, 2015).

Chapter 6 uses the validated partial equilibrium model described in Chapter 5 to analyze policy pathways outlined in the US Environmental protection Agency (EPA) Clean Power Plan (CPP). A baseline projection to 2030 shows that coal capacity retirements, combined with fuel switching from coal to gas power in response to low gas prices and continued expansion in renewables, lead to a CO₂ emission reduction of almost 28% in the US electricity sector. This means that the 32% reduction goal of *the CPP can nearly be met without any policy intervention*. Still, there is a need for some small interventions to meet the stated goal. The CPP describes three building blocks, one of which focus on capacity utilization (building block 2) and another on capacity expansion (building block 3). Therefore, the model detailed in Chapter 5 is ideally suited for analyzing policies given the CPP's building block framework.

Two policy alternatives, a carbon tax and an additional investment subsidy for wind and solar, are presented and compared to a baseline for 2030. The carbon tax reduces CO₂ by fuel-switching from coal to gas (building block 2) and by expanding renewables due to reduced returns to fossil-fuel capacity (building block

3). On the other hand, wind and solar investment subsidies increase returns for renewable (specifically wind and solar) capacity (building block 3), but the capacity expansion crowds-out utilization and actually increases emissions from fuel-switching (building block 2) as compared to the baseline. This is an important consideration for pursuing an investment subsidy policy as opposed to the economically-efficient, but less politically tractable carbon tax. We also observe that wind and solar investment subsidies crowd-out returns to capacity for other renewable options (e.g. nuclear, hydro, geothermal, biomass). Finally, because the investment subsidy case does not penalize returns to coal capacity (the largest contributor to emissions), coal capacity retires at a slower rate. These insights provide specific guidance in designing policies within the CPP policy framework.

Chapter 7 reviews the advances in database construction (Chapters 2–4) and the representation of the electricity sector in computational equilibrium models (Chapter 5) and discusses the insights uncovered in the analysis of the US EPA CPP (Chapter 6). The chapter then describes two future steps. First, the computational model can be integrated into a CGE framework to analyze the CPP in terms of welfare and other economy-wide impacts. Second, the role of the US electricity sector as part of the global economy can be studied by integrating regional trade flows to study important global trends like international carbon policy and possibilities for US energy exports.

CHAPTER 2. MATRIX BALANCING WITH UNKNOWN TOTAL COSTS: A NOVEL METHOD FOR GTAP-POWER

Global economic analysis, specifically computable general equilibrium, relies on an underlying database that describes the inter-sectoral and inter-regional linkages in an economy. This type of database is typically constructed from national input–output tables, trade, macroeconomic, and trade protection data from several sources. However, due to the aggregate nature of the data, the sectoral resolution tends to be quite coarse. For instance, the GTAP database, which is the predominant database for global CGE analysis, has only one aggregate electricity sector which encompasses production, collection, and distribution. The database identifies fuel inputs, but is blind to the actual technologies used to produce electricity. It is particularly non-informative for non-fuel based technologies (e.g. nuclear, hydroelectric, wind, solar), which encompass roughly 32% of the global electricity sector. Technology-specific advancements and policies motivate a more detailed representation of electricity generation in CGE databases.

In the GTAP-Power database, described in the Chapters 2–4, the aggregate electricity sector in GTAP is disaggregated into the following new sectors: transmission and distribution (T&D), nuclear, coal, gas, oil, hydroelectric, wind, solar, and other power technologies. Gas, oil, and hydroelectric power are further differentiated by load type: base and peak (henceforth designated by “BL” and “P” suffixes, respectively).¹ The intent of the split between base and peak load is two-fold. First, total generation data comes in the form of fuel inputs (e.g. GWh generated from natural gas); however, several different technologies (e.g. combustion turbine, steam turbine, combined-cycle) convert the fuels into electricity. These technologies

¹Chapters 2 and 3 show matrix balancing methods and their implications on modeling results using disaggregations with less technological detail for clarity.

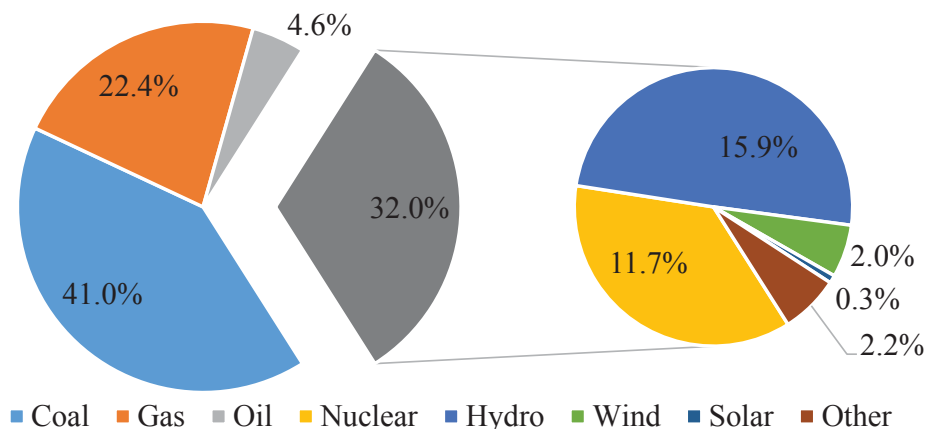


Figure 2.1.: Shares of global electricity generation by technology. About 32% comes from non-fuel-based technologies and would be represented as a portion of total capital in an aggregate electricity sector.

have cost structures that must be differentiated, and in a typical case where a modeler wishes to re-aggregate these into a single fuel-based sector (e.g. gas power) using this database this is important.² Second, connecting the data to modeling, base and peak load are distinct types of generation. Without differentiating electricity production by these operational considerations, a model can have a technology like solar taking over the entire generation, which is not realistic in the modern electricity system (i.e. without storage for time arbitrage).

The objective of the electricity sector disaggregation is to include as much engineering-level data as possible and preserve the economic relationships that they imply subject to the constraint imposed by the aggregate electricity sector. Chapters 2–4 all make individual, but related contributions to the construction of an electricity-detailed computable general equilibrium database. The relevant data are reviewed in Section 2.1 and the remainder of the chapter introduces a novel matrix balancing method that best preserves important economic relationships (i.e. relative

²In the long-run specific technologies such as combined-cycle, combustion turbine, and steam turbine gas would provide a better idea of costs, but the modeling issues of how each of these technologies compete from an operational perspective is still unclear. Therefore, a simple aggregate base and peak load differentiation balances operational considerations and data availability.

contributions to input demand across technologies (termed “row share”) and cost structure). Later Chapter 3 shows that these economic relationships translate directly into modeling results and reinforce the strength of the matrix balancing method in the final GTAP-Power database. Specific details of the full GTAP-Power disaggregation are highlighted in Chapter 4 along with some summary results.

2.1 Engineering-level data⁴

The foundations of the bottom-up economic data are: i) electricity production (in GWh) by technology (IEA, 2015, 2014) and ii) levelized capital, fuel, operating and maintenance (O&M) costs for each technology (IEA/NEA, 2010). Levelized costs of electricity (LCOE) are annualized unit costs (US\$ per GWh).⁷ Electricity production is represented as the matrix \mathbf{Q}^g with elements q_t^g where t is the index for the new sector. Levelized costs are represented as the matrix \mathbf{L}^0 with elements l_{it}^0 where i is the index for the input cost. The super-script 0 indicates that this is reported data. The final levelized costs, \mathbf{L} , are the balanced costs. The base years for this disaggregation are 2004, 2007, and 2011. These data sources are constructed from the reported data along with some elementary assumptions – e.g. missing costs are filled in the manner outlined in Peters (2015).

It is worth noting that levelized costs of electricity are notoriously misleading estimates of the value of electricity. First, each individual levelized cost is derived from a number assumptions that are unobservable *ex ante* (e.g. depreciation rate,

⁷The use of levelized capital costs requires some caveat for general usage in disaggregations of the electricity sector. Additional care and manipulation may be required to capture the financial structure of the sector based on the research question at hand. For instance, a large portion of the value in capital costs are construction costs which likely have taken place in previous years. This can be problematic for some I–O analysis and especially life-cycle analysis where construction embodied in levelized capital costs are excluded (Marriott, 2007; Lindner et al., 2013). Methods in Marriott (2007) can be used to create an initial matrix A for construction-specific costs rather than capital in the application presented here. Levelized capital costs are useful in disaggregating social accounting matrices and computable general equilibrium databases. The application here is for GTAP, a SAM/CGE database, so we use levelized capital costs to allocate capital in the database.

Table 2.1.: Total electricity production by technology in the United States in 2011, Q^g (in TWh) and shares. Results may not sum due to rounding. Source: IEA (2014, 2015); Peters (2015).

	Nuclear	Coal	GasBL	GasP	Oil	Hydro	Wind	Solar	Other	Total
TWh	821.4	1,872.2	445.1	611.4	31.4	321.7	120.8	6.2	96.3	4,326.6
Share	19.0%	43.3%	10.3%	14.1%	0.7%	7.4%	2.8%	0.1%	2.2%	100%

Table 2.2.: Levelized cost for electricity production by technology in the United States, l_{it}^0 (2011 US\$/MWh). Source: IEA (2014, 2015); Peters (2015).

LCOE	Nuclear	Coal	GasBL	GasP	Oil	Hydro	Wind	Solar	Other
Capital	27.4	19.7	9.2	5.9	11.3	89.3	60.4	202.9	23.5
O&M	22.9	8.8	3.7	4.6	22.8	13.3	16.6	17.2	20.2
Coal	-	20.2	-	-	-	-	-	-	-
Gas	-	-	50.8	68.6	-	-	-	-	-
Oil	-	-	-	-	214	-	-	-	32.2

project lifetime). Second, the value of electricity varies over time, space, and contract type (Joskow, 2011; Hirth et al., 2014).

Despite the deficiencies of levelized costs in estimating the value of power generation, Table 2.2 shows how they can capture important economic differences between the technologies. For example, wind, hydroelectric, and solar power have no specific fuels costs, but have high capital costs. As such, their costs are not sensitive to fuel shocks, but are sensitive to the cost of capital, which motivates support for renewables via investment tax credits in the United States.

There are only limited input cost data, but the I–O component of the GTAP database comprises 63 different input costs to the electricity sector. The vector of 63 inputs is aggregated to map to available aggregate input cost categories (capital, fuels, and O&M) to create the row constraints, u_i . This results in the row constraint for aggregate inputs to the new electricity industries where each row sums to u_i .

The matrix of initial estimates, \mathbf{A} , can be constructed from the economic data. The estimates for a_{it} are as follows:

$$a_{it} = \frac{l_{it}^0 \cdot q_t^g}{\sum_i \sum_t l_{it}^0 \cdot q_t^g} \cdot \sum_i u_i \quad (2.1)$$

where l_{it}^0 is the levelized cost for cost i for technology t , and q_t^g is the GWh production for technology t . The bottom-up costs are normalized to the total value in the original electricity sector (i.e. $\sum_i u_i$).

Table 2.3.: Target matrix, \mathbf{A} , for the United States (in millions of 2011 US\$) derived from levelized cost and production data compared to the GTAP totals (\mathbf{U}).

	T&D	Nuc.	Coal	GasBL	GasP	Oil	Hydro	Wind	Solar	Other	Total	U
Cap.	24,770	22,481	36,880	4,101	3,627	356	28,736	7,303	1,248	2,263	131,763	129,967
O&M	43,676	18,811	16,542	1,658	2,826	716	4,286	2,011	106	1,944	92,575	133,139
Coal	-	37,884	-	-	-	-	-	-	-	-	37,884	61,781
Gas	-	-	-	22,625	41,957	-	-	-	-	-	64,582	48,361
Oil	16	-	-	-	-	6,724	-	-	-	3,103	9,843	7,758

The final two columns of Table 2.3 show that the totals implied by the bottom-up data do not match that of the corresponding costs in the original GTAP electricity sector.⁸ Therefore, the target matrix \mathbf{A} must be balanced to meet the accounting consistency constraints in the GTAP database, \mathbf{U} . The remainder of this chapter details a novel matrix balancing method specifically designed for the electricity disaggregation problem and available data.⁹ Similar to bottom-up and top-down

⁸Levelized costs try to capture the annualized cost of production while considering data such as overnight capital costs, depreciation rate, fuel costs averaged over the year, heat rate, and other technological factors. Top-down data comes from the reporting of final reported costs in broad categories which may encompass a wider breadth of costs than the bottom-up data (e.g. transmission and distribution, insurance services, customer service, litigation). These different perspectives offers some insight on the origin of the discrepancies between the two data types.

⁹Both Chapters 2 and 3 use GTAPv8 data with a base year of 2007 and an alternate scheme that only includes the following technologies: nuclear, coal, gas, oil, hydroelectric, wind, and solar. The intent is for clearer analysis of the properties of the matrix balancing techniques shown and discussed in Peters and Hertel (2015b,a). The full estimates shown in Table 2.3 are used for the full GTAP-Power disaggregation in Chapter 4. The data is otherwise identical.

modeling discussed in Chapter 1, levelized cost (\mathbf{L}^0) and production quantities (\mathbf{Q}^g) are bottom-up data that must be balanced to conform to fit the top-down totals (i.e. \mathbf{U}). In this way, the gulf between the two modeling philosophies is drawn slightly closer together.

2.2 Matrix Balancing with Unknown Total Costs

Systems of national accounts (e.g. supply and use tables, input–output (I–O) tables, social accounting matrices (SAM), and computable general equilibrium (CGE) models) are commonly used for multi-sectoral, multi-regional economic analysis. The I–O table comprises the core of industry production information within SAM and CGE databases, yet it often has insufficient detail for addressing specific issues - for example, regulating CO₂ emissions from the electric power sector. This section on the disaggregation of an I–O industry in the unique situation where information on total costs of the new sectors are either unknown or deemed less “trustworthy” than the component costs (e.g. levelized input costs for each technology).

In the disaggregation problem, detailed sub-sector information must be balanced with the aggregate sector using a matrix balancing method. In a survey of these methods, Huang et al. (2008) recommends RAS (a biproportionate cross-entropy method)¹⁰ and the improved normalized squared differences (INSD) method (Friedlander, 1961) for balancing I–O tables. Similarly, Temurshoev et al. (2011) analyzed prevailing methods for projecting supply and use tables and concluded that RAS, INSD, as well as Kuroda (1988) perform better than the alternatives explored in these works.¹¹

However, Kuroda’s method (with some minor modification discussed later) is the only one of these capable of addressing the case where total costs are unknown and/or where the researcher does not wish to impose fixed conditions on the total costs.

¹⁰RAS has many variants to improve upon its various limitations, notably GRAS (Lenzen et al., 2007). The reference to “RAS” in this work is general to the variants as they are basic cross-entropy formulation. Our example only considers a positive matrix, so GRAS is not required.

¹¹Kuroda’s method was not included in Huang et al. (2008).

With no hard constraint on total costs, both RAS and INSD reduce to the pro rata distribution which only considers relative contributions to input demand and ignores cost structure of the technologies.

While RAS and INSD target the individual elements, Kuroda’s method targets the economic relationships of cost structure and relative contributions to input demand.¹² As will be shown, both of these economic relationships are critical in determining the economic outcomes of studies involving policies bearing on the sector being disaggregated. The cost structure is of importance in the face of an input (e.g. fuel) price shock. The row share has direct implications in terms of modeling substitution between the disaggregate sub-sectors. If the ordering of relative input shares in the reported data (e.g. capital share to each sub-sector) is not preserved, the substitution between the technologies may be distorted and even reversed (McDougall, 1999). These two economic relationships are non-trivial when the database is implemented in even the simplest of models. They also play a key role in designing an objective to balance a matrix with unknown total costs.

In this section, we slightly adjust the Kuroda method so the column constraint (termed Kuroda-NC) can be removed, and we propose the share preserving cross-entropy (SPCE) formulation which attempts to preserve cost structures and row shares without any constraint on total cost. The cross-entropy (CE) formulation complements the sum squared error (SSE)-type approach of Kuroda.¹³ We also show that the SPCE solution reduces to the RAS solution when a total cost constraint is imposed.

The motivating example for this research on unknown total costs is found in the United States where policymakers have turned to regulations, taxes, and subsidies on specific electric power sub-sectors to limit CO₂ emissions. An aggregate electric power

¹²Henceforth, we will refer to cost structure interchangeably with “column shares.” The relative contributions to input demand across industries will be referred to as “row shares.” The latter refers to the share of input use by a given sub-sector in overall input use (e.g. capital employment in different electricity technologies). Alternatively, the row share can be thought of as the production-weighted input intensity across industries or “sales share.”

¹³It is important to note that we do not wish to determine the virtues between cross-entropy, sum squared error, or other objectives. These largely remain the preference of the individual researcher.

sector, as is typically found in I–O tables, is insufficient for capturing the effects of technology-level policies where different technologies have significantly different cost structures and substitute for one another in the ultimate production of electricity. Total costs of electricity production from different technologies are largely unknown, and the price of electricity differs across generating technologies due to operational realities (e.g. “base” versus “peak” power). However, there exist economic data regarding the generating technologies (i.e. cost structure and row share) which can inform how these technologies may respond in the face of technological, economic, or policy shocks. Despite this, previous work on electricity disaggregation has not focused specifically on these dual relationships (Lindner et al. (2013); a detailed discussion is presented later). In this example, the relative component costs are deemed more “trustworthy” than the summation. The lack of proper total cost data renders most existing formulations impotent, with the exception of SPCE and a variant of Kuroda’s method.

While the electricity sector motivates the disaggregation problem which seeks to preserve economic relationships where total costs are unknown, the proposed solution can be applied to other aggregate sectors which employ different technologies to produce slightly dissimilar end products preventing a simple assignment of a uniform price across the technologies (e.g. services, biofuels). In addition, the methods discussed in this work are also not exclusive to the sector disaggregation problem and can be used broadly across other matrix balancing problems.

This section is organized as follows. The structure of the disaggregation problem and important features of the electricity application are outlined in Section 2.3. The importance of cost structure and row share in application are detailed with a simple set of equations in Section 2.4. Next, the modified Kuroda-NC and the novel share preserving cross-entropy approach that is well-suited for the unknown column sum problem, termed SPCE, are formulated. The SPCE objective returns the RAS solution when a column sum constraint is imposed and should outperform RAS when the constraint is removed because the constraint becomes unnecessarily

restrictive. Section 2.6 analyzes RAS, SPCE, Kuroda, and Kuroda-NC with a practical application to the electric power generation sector using a simplified disaggregation of the 126 regions in GTAPv8 and compares their performance with respect to various performance indicators. A comparison between RAS and SPCE and between Kuroda and Kuroda-NC shows how performance can be improved by relaxing the total cost constraint. Comparing SPCE and Kuroda-NC shows how the two different definitions of closeness (entropy versus SSE, respectively) perform.

2.3 The Structure of the Disaggregation Problem

I–O industries are often characterized by aggregate sectors due to the lack of specific sub-sector information. However, policies and technological advances are often related to specific technologies/sub-sectors and may not be applicable to the aggregate sector. Disaggregating the aggregate industry in I–O tables is an important step in reconciling bottom-up technological detail with top-down economy-wide modeling (Sue Wing, 2008; Lenzen, 2011). The disaggregation problem presented here focuses on how to transition from an aggregate sector to several sub-sectors when the total costs of the new industries are unknown - all while preserving important economic relationships within and across the sub-sectors.

2.3.1 The Disaggregated I–O table

The fully disaggregated supply-side matrix is constructed by disaggregating a particular sector (e.g. electricity) into sub-sectors while the other sectors remain unaffected. The balanced disaggregation is defined as matrix \mathbf{X} with elements x_{it} where i is an input in the same vector of inputs as those in the full GTAP database, and t is a new industry within the set of new industries (or technologies) which are inserted in place of the aggregated sector. By way of example, x_{it} might refer to capital inputs into the solar power generation sub-sector. In order to perform this

	Technologies (t)					Sum ($x_{i\bullet}, u_i$)
Inputs (i)						
Sum ($x_{\bullet t}, v_t$)						

Figure 2.2.: The I–O table for the disaggregated industry (\mathbf{X})

disaggregation, an initial matrix \mathbf{A} is constructed from economic and/or technological information about alternative technologies (e.g. Equation 2.1 in the previous section).

The disaggregation problem is to minimize the distance between the elements of \mathbf{X} and \mathbf{A} subject to a set of constraints imposed by the I–O structure (Schneider and Zenios, 1990). In particular, the sum of x_{it} over all t (row sum) must equal the original employment of input i in the aggregate sector defined as u_i (i.e. $x_{i\bullet} = u_i$ where $x_{i\bullet} \equiv \sum_t x_{it}$).¹⁴ The majority of methods also impose a column sum constraint on the sum of x_{it} over all i for each t must equal some given value v_t (i.e. $x_{\bullet t} = v_t$ where $x_{\bullet t} \equiv \sum_i x_{it}$). However, this constraint is not strictly required for consistency since the earlier row sum restriction will ensure that total value in the disaggregate matrix will equal that of the aggregate industry. Therefore, it is preferable to avoid a potentially restrictive column sum constraint when information on the column sum, v_t , is unknown or of less reliability than the component costs (reinforced later in Chapter 3).

The disaggregated industry matrix (\mathbf{X}) replaces the aggregate industry in the full matrix to construct a complete GTAP-Power database containing the new disaggregated electricity industries along with those in the GTAP database.

¹⁴Throughout this work, summations are assumed to range over the entirety of the dimensions unless otherwise stated.

2.3.2 The Nature of Economic Data for an Electricity Disaggregation

As discussed in Section 2.1, the fundamental economic data which can be reasonably be obtained are: i) data on total generation from each technology in gigawatt-hours (GWh), ii) the input costs to each generation technology (US\$ per GWh), and iii) the original employment of inputs in all electricity generation, u_i (in US\$). In this case, available input cost data consist of capital, fuel, and operating and maintenance (O&M) levelized cost (i.e. annualized cost per GWh) data reported with different error distributions for each category of input. Total cost data for generating technologies are not available because national accounts do not generally report on detailed technologies. Because national accounts are constructed from plant-level data, sub-sector total cost data exist in theory; however, in practice, depending on the region, this type of data can be proprietary, unreliable, and unpublished. Constructing detailed national accounts from plant-level data would obviously overcome the unknown total cost problem.

Matrix balancing approaches would normally advise to construct artificial column totals using the available data. One way to construct total costs of generation is to impute a uniform price of electricity from the original values and GWh produced (Lindner et al., 2012; Shrestha and Marpaung, 2006); however, in reality unit costs of electricity differ, sometimes greatly, between generation types due to operational realities (e.g. “base” versus “peak” power, vintage of the power plants, etc.).

Alternatively, assuming the cost data are complete (which is often not the case), column sums could be derived from summing component costs for a technology and normalizing to fit the total value in the original I–O data.¹⁵ However, fundamentally different methods of measurement and assumptions are used to construct the levelized

¹⁵Robinson et al. (2001) highlights the importance of known column sums in CE, and offers the use of weighted errors in total costs as a powerful tool where column sums are uncertain. As proposed, simple summing up the reported input costs has implications for the error distribution of the column sum when the individual cost errors are dependent in the cost structure or across industries (which is likely in the case of data from the same source). This complicates a simple “cost totals with error” measure, especially in the case with several weighted error on various constraints in the model. The intent here is to avoid the additional unnecessary constraint altogether.

input costs than are used to create the aggregate national accounts.¹⁶ Thus, this approach unnecessarily imposes a total cost constraint that is derived from two or more disparate data sources. Furthermore, the development of these column constraints also requires complete knowledge of input costs, which may not be available.

It is more intuitive, and most general, to preserve the economic relationships in the input data rather than construct total cost data from incomparable sources solely to fit into an existing method, especially when the presence of a constraint unnecessarily restricts the result.

2.4 The Importance of Preserving Economic Relationships in Modeling

The two economic relationships we consider here are cost structure and row share. Both appear often in economic modeling using I–O, SAM, and CGE-type analysis. Cost structure and row share are defined, respectively, as:

$$c_{it} \equiv \frac{x_{it}}{x_{\bullet t}} \quad \text{and} \quad r_{it} \equiv \frac{x_{it}}{x_{i\bullet}} \quad (2.2)$$

To illustrate their importance, consider Equations 2.3–2.8 which provide a linearization of a generic, competitive, long-run partial equilibrium model (in the spirit of ORANI (Dixon, 1982)) with two activities (A and B) which substitute imperfectly for one another in the supply of electricity to the grid, with elasticity of substitution in use, σ . These sectors could be nuclear-powered (high capital cost) and gas-powered electricity generation (lower capital cost) where the aggregate output of electricity is produced by substituting between nuclear and gas-power. This local approximation to the percentage change in demand for each type of power generation may be written as a function of the percentage change in total electricity production,

¹⁶Levelized costs are annualized costs of production which consider data such as overnight capital costs, depreciation rate, fuel costs, heat rate, and other technological factors. National accounting data comes from the reporting of final reported costs in broad categories which may encompass a wider breadth of costs (e.g. transmission and distribution, insurance services, customer service, litigation).

\widehat{q}_{ely} , as well as the change in price of each individual type of power ($\widehat{p}_A, \widehat{p}_B$), relative to the change in the average cost of electricity, \widehat{p}_{ely} :

$$\widehat{q}_A = \widehat{q}_{ely} - \sigma \cdot (\widehat{p}_A - \widehat{q}_{ely}) \quad (2.3)$$

$$\widehat{q}_B = \widehat{q}_{ely} - \sigma \cdot (\widehat{p}_B - \widehat{q}_{ely}) \quad (2.4)$$

When the two types of power are highly substitutable ($\sigma \gg 0$), large swings in the source of power generation can be expected. When they substitute poorly in use, we can expect nuclear and gas-powered generation to rise and fall, more or less in tandem with changes in total power demand, \widehat{q}_{ely} . Assuming either average cost regulation or competitive entry and exit of generating technologies in the long run, we can derive a zero profit condition for each sub-sector, as well as for the industry as a whole. These equations simply state that the sub-sector output price change in long-run equilibrium will be determined by the cost-share weighted sum of input price changes.

$$\widehat{p}_A = \sum_i c_{iA} \cdot \widehat{p}_{iA} \quad (2.5)$$

$$\widehat{p}_B = \sum_i c_{iB} \cdot \widehat{p}_{iB} \quad (2.6)$$

$$\widehat{p}_{ely} = \theta_A \cdot \widehat{p}_A + (1 - \theta_A) \cdot \widehat{p}_B \quad (2.7)$$

where \widehat{p}_{it} is the percentage change in the price of input i used in technology t and θ_t is the share of technology t in total electricity cost. We will assume for the time being that the price change of input i is identical across technologies (i.e. $\widehat{p}_{iA} = \widehat{p}_{iB}$).

From the zero-profit conditions, Equations 2.5 and 2.6, it can be seen that the cost share of an input for a given activity will have important consequences for the power sector's long run response to a change in the price of that input. Consider, for example, the case in which the price of a common input, say capital, rises by 20%. The intensity of input i in each productive sector t , (i.e. column coefficient in the I-O context, c_{it}), determines the size of the two relative price changes for

activities A and B , with the relative price of the more capital intensive sector (let us assume this is A) rising more. Faced by a relative price rise for activity A , the power sector will substitute toward activity B so that $\hat{q}_A < 0$ and $\hat{q}_B > 0$. However, if, in the process of disaggregating the power sector, the relative capital intensity of activity A is not preserved, i.e. now, $c_{iB} > c_{iA}$, then the direction of quantity change is reversed when the cost of capital rises. This is very problematic in cases where the ordering of column shares is not necessarily preserved. In fact, the minimum sum of column cross-entropy (MSCCE) posed by Golan et al. (1994) and studied by Robinson et al. (2001) has been shown to flip the ordering of input intensity across sectors McDougall (1999), which highlights a danger in preserving only one of these economic relationships. Simply removing the column constraint in the RAS and INSD formulations reduces both methods to the pro rata distribution which preserves only the row share.

Now turn to the related question of row shares. For the sake of simplicity, we will assume that the input demands in the individual activities change in fixed proportion to output levels (e.g. no capital-labor substitution in the individual technologies used to generate power). We then can add one more equation to the model which serves to determine the change in total use of input i (e.g. capital), \hat{q}_i , in the aggregate electricity industry. This can be written as the quantity share-weighted sum of sub-sector output changes. Since both sectors face the same input prices, the quantity and value-shares are identical, so that:

$$\hat{q}_i = r_{iA} \cdot \hat{q}_A + r_{iB} \cdot \hat{q}_B \quad (2.8)$$

Recall r_{iA} represents the row share of sector A in the total use of input i . From this, it is easy to see that if sector A (e.g. nuclear power) is relatively more capital intensive than B (gas generation), then a shift from nuclear to gas will lessen the demand for capital in the power sector. Reversing this row share relationship can result in misleading estimates of the total input requirements for the power sector following an exogenous price shock or policy intervention.

Clearly, both the cost structure and relative contributions to input demand are important when it comes to economic analysis of a sector comprising varied individual technologies. Targeting these economic relationships, rather than the values themselves, can allow the relaxation of a total cost constraint. This is illustrated further later in this dissertation.

2.5 Possible Approaches for Unknown Total Costs

Huang et al. (2008) and Temurshoev et al. (2011) compare various matrix balancing methods and reach similar conclusions: RAS (e.g. Lenzen et al. (2007)), INSD (Friedlander, 1961), and Kuroda (1988) are preferred methods. Therefore, we confine our analysis to these three alternatives, although there is a large and diverse literature on matrix balancing and other reasonable alternatives might exist.

The motivation for this research is in the case of unknown total costs or when the total cost constraint is thought to be unreliable compared to other aspects of the initial data used to construct the matrix \mathbf{A} . If the corresponding constraint is removed from RAS and INSD formulations, they both will reduce to a pro rata distribution which only considers row share and ignores cost structure. This leaves us with Kuroda as the only preferred option capable of the case with unknown or unreliable total costs. The Kuroda objective directly targets cost structure and row shares so that the total cost constraint can be removed without compromising the intent of the method.

However, many researchers prefer the information-theoretic cross-entropy approach (see Shannon (1948) and Kullback and Leibler (1951) for an in-depth information-theoretic treatment). To the authors' knowledge there is no such approach which allows the flexibility of removing the total cost constraint without adversely impacting the biproportional intent of the method. We therefore propose such a novel entropy-type approach to complement the sum squared error (SSE)-type Kuroda approach for the case of unknown or unreliable total costs.

Accordingly, the Kuroda, Kuroda without a total cost constraint (Kuroda-NC), and the novel share preserving cross-entropy (SPCE) are formally introduced as constrained optimization problems in this section.

2.5.1 Kuroda's Method without a Total Cost Constraint (Kuroda-NC)

Temurshoev et al. (2011) reviews three different weighting schemes for the original Kuroda objective function, but the original Kuroda (1988) weighting scheme of “equal percentage change” is determined to perform best. The objective function is

$$\min_{x_{it}} \frac{1}{2} \sum_i \sum_t \left[\left(\frac{x_{it}/u_i}{r_{it}^0} - 1 \right)^2 + \left(\frac{x_{it}/v_t}{c_{it}^0} - 1 \right)^2 \right] \quad (2.9)$$

where

$$c_{it}^0 \equiv \frac{a_{it}}{a_{\bullet t}} \quad \text{and} \quad r_{it}^0 \equiv \frac{a_{it}}{a_{i\bullet}} \quad (2.10)$$

which correspond to the row share and the cost structure of the target matrix \mathbf{A} , respectively. Elements u_i and v_t are the exogenous target row and column totals for the balanced matrix, \mathbf{X} , respectively. The objective essentially minimizes the sum squared percentage error in both row share and cost structure. For matrix balancing this objective is typically subject to both row and column constraints.

$$\sum_t x_{it} = u_i \quad \text{for all } i \quad (2.11)$$

$$\sum_i x_{it} = v_t \quad \text{for all } t \quad (2.12)$$

So that the Kuroda method is described by Equations 2.9, 2.10, and 2.11. To relax the total cost constraint the necessary modification to the Kuroda objective is straightforward. The Kuroda-NC objective can be written as:

$$\min_{x_{it}} \frac{1}{2} \sum_i \sum_t \left[\left(\frac{x_{it}/u_i}{r_{it}^0} - 1 \right)^2 + \left(\frac{x_{it}/x_{\bullet t}}{c_{it}^0} - 1 \right)^2 \right] \quad (2.13)$$

where the column sum in the balanced matrix, \mathbf{X} , is no longer exogenously given by v_t , but rather endogenously determined by $x_{\bullet t}$. Thus, the Kuroda-NC method can be described by Equations 2.11 and 2.13. The result is identical to the Kuroda result when the column constraint, Equation 2.12, is also included.

2.5.2 RAS and Share Preserving Cross-Entropy (SPCE)

To the authors' knowledge there is no previously documented entropy-theoretic approach which is capable of relaxing the total cost constraint (i.e. Equation 2.11) while preserving both row share and cost structure. For example, the commonly-used RAS approach reduces to the pro rata distribution when the column constraint is removed.

Therefore, this section presents the share-preserving cross-entropy (SPCE) objective - an alternative entropy objective function specifically designed to preserve economic relationships (i.e. cost structure and row share) in the face of unknown column sums but known component costs. The component costs might represent economic aggregates for which data are available (e.g. operating and maintenance costs, labor, fuels). The result can then provide estimates of column sums of each of these economic aggregates which can subsequently be used with traditional CE methods, like RAS, to derive values from sub-components (e.g. unskilled and skilled labor costs derived from an aggregate labor cost) while still preserving the economic relationships in the original data as best as possible.

The interpretation of cross-entropy is the divergence between the *a priori* (unbalanced) and the *a posteriori* (balanced) distribution. The RAS objective in its simplest form is as follows

$$\min_{x_{it}} \sum_i \sum_t x_{it} \cdot \ln \frac{x_{it}}{e \cdot a_{it}} \quad (2.14)$$

where e is the base of the natural log.¹⁷ The RAS method encompasses Equation 2.14 subject to both the constraints given in Equations 2.11 and 2.12.

Cross-entropy methods, such as RAS, are built on information-theoretic foundations. Kullback and Leibler (1951) denote the mean information to discriminate between two hypotheses given some observation; the discrete distribution version of which, applied to our two-dimensional problem, is given by

$$D(p : q) = \sum_i \sum_t \left(\frac{x_{it}}{x_{\bullet\bullet}} \right) \cdot \ln \left(\frac{x_{it}/x_{\bullet\bullet}}{a_{it}/a_{\bullet\bullet}} \right) \quad (2.15)$$

where $x_{it}/x_{\bullet\bullet}$ is the *a posteriori* distribution and $a_{it}/a_{\bullet\bullet}$ is the *a priori* distribution of values. Given that $a_{\bullet\bullet}$ is non-negative and constant (i.e. exogenous) and some constraint imposes $x_{\bullet\bullet}$ to be constant as well (e.g. Equation 2.11 or 2.12), it is straightforward to show that minimizing cross entropy is equivalent to minimizing the Kullback and Leibler (1951) measure of information needed to discriminate between two discrete distributions, although the objectives are not identical.

In information-theoretic terms, minimizing cross-entropy is equivalent to minimizing the entropy information gained (i.e. extraneous information) from imposing the balancing constraints on the *a priori* distribution. Alternatively, Junius and Oosterhaven (2003) interpret this as the information loss from the *a priori* distribution due to the matrix balancing process. The interpretation of the RAS objective in Equation (2.14) considers the elements of matrix \mathbf{A} to be the relevant *a priori* distribution.

Instead the SPCE objective considers both cost structure and row share in the matrix \mathbf{A} to be the *a priori* distributions. As such the objective is defined as follows

$$\min_{x_{it}} \left\{ \sum_i \alpha_i \left[\sum_t \frac{x_{it}}{x_{i\bullet}} \ln \left(\frac{x_{it}/x_{i\bullet}}{r_{it}^0} \right) \right] + \sum_t \beta_t \left[\sum_i \frac{x_{it}}{x_{\bullet t}} \ln \left(\frac{x_{it}/x_{\bullet t}}{c_{it}^0} \right) \right] \right\} \quad (2.16)$$

¹⁷The RAS objective function is sometimes written without the base of the natural log shown in Equation 2.14. Subject to any constraint that fixes $x_{\bullet\bullet}$ (e.g. Equations 2.11 and 2.12), the minimization of the RAS objective function with and without the base of the natural log are equivalent.

where α_i and β_t are arbitrary weights on the row share cross-entropy and the cost structure cross-entropy, respectively. With an endogenous weighting scheme of $\alpha = x_{i\bullet}$ and $\beta = x_{\bullet t}$, the SPCE objective becomes

$$\min_{x_{it}} \sum_i \sum_t x_{it} \cdot \ln \left(\frac{x_{it}/x_{i\bullet}}{r_{it}^0} \cdot \frac{x_{it}/x_{\bullet t}}{c_{it}^0} \right) \quad (2.17)$$

Appendix A in Peters and Hertel (2015b) discusses the first order conditions and a method performance check to show that imposing the row and/or column totals directly from \mathbf{A} as constraints returns the full matrix \mathbf{A} . The column constraint can be relaxed while still preserving the important economic relationships, and the SPCE would then comprise Equations 2.11 and 2.17. This objective also allows the row constraint to be relaxed instead of the column constraint (i.e. Equations 2.12 and 2.17); however, the practical use of this is unclear.

Another important feature of this objective formulation is that if the column constraint is also included (i.e. Equations 2.11, 2.12, and 2.17) the result is identical to the RAS solution (see Appendix B in Peters and Hertel (2015b)). This makes the solution especially attractive for comparison purposes and for practitioners who are most confident in the RAS solution.

There may be some concern regarding the endogenous weighting scheme which leads to Equation 2.17. It is important to note that if the row total constraint is imposed (Equation 2.11) then α_i is in fact exogenously given by u_i , and if the total column constraint is imposed (Equation 2.12) then β_t is exogenously given by v_t . However, in the context of unknown total costs we assume that the total column constraint is not imposed and β_t is in fact endogenous and may lead to some circularity. Thus, we compare the SPCE (Equations 2.17 and 2.11) result with $\beta_t = x_{\bullet t}$ (endogenous weighting) against the RAS results as well as two alternate exogenous weighting schemes in Equation 2.16: $\beta_t = a_{\bullet t} (\sum_i u_i / a_{\bullet\bullet})$, termed SPCE-1, and the uniform weighting scheme $\beta_t = 1/m \sum_i \alpha_i$, termed SPCE-2. Figure 2.3 shows a frequency chart of the absolute percentage deviation of x_{it} from a_{it} for capital and

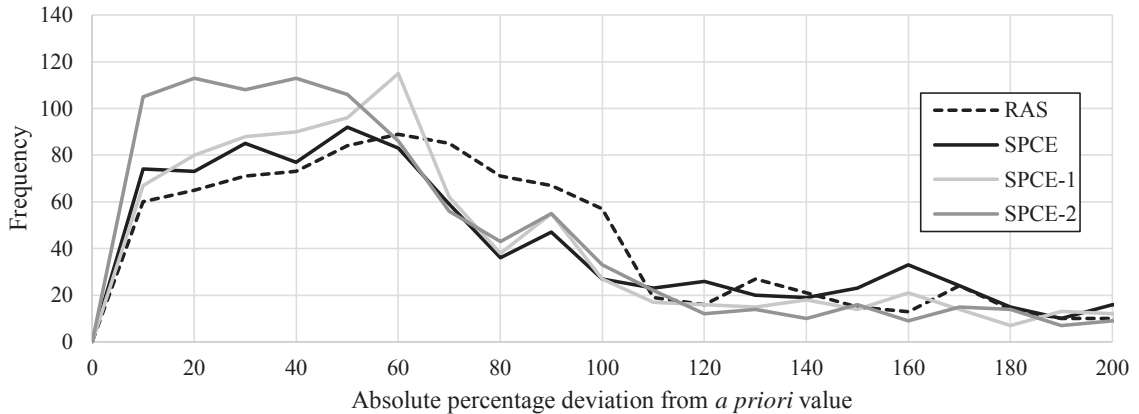


Figure 2.3.: A frequency chart of absolute percentage deviation of the balanced values from the *a priori* values show that the endogenous weighting scheme of the SPCE formulation is comparable to other exogenous weighting schemes.

O&M across each technology and region. We see the deviations of the endogenous weighting scheme (SPCE) are not significantly different than RAS or the exogenous weighting schemes. From this, we can conclude that the endogenous weighting appears robust in most cases.

The following section compares the Kuroda, Kuroda-NC, RAS, and SPCE formulations with an application to the electricity sector where component costs are deemed more reliable than their totals. In other words, the preservation of economic relationships between component costs is deemed more important than strictly adhering to the implied total costs.

2.6 An Example in a Simple Electricity Sector Disaggregation

The motivating application of this specific formulation is the disaggregation of an aggregate I–O electricity sector into sub-sectors which include distinct generating technologies. The I–O table itself can be used for a number of purposes (e.g. understanding economic structure, I–O analysis, life-cycle analysis, CGE modeling). Electricity technologies are highly substitutable. Therefore, as reinforced later in Chapter 3, preserving the associated economic relationships is of utmost importance

when undertaking economic analysis of policy interventions. Failing to preserve the economic relationships should be seen as a serious limitation in the context of electricity sector disaggregation efforts meant to include generating technologies.

One major criticism of previous work aimed at power sector disaggregation is that the I–O construction methodology is not always transparent, perhaps due to *ad hoc* methods (e.g. Han et al. (2004); Limmeechokchai and Suksuntornsiri (2007)). This is especially true in large-scale CGE and IAM research where documentation is copious and technical details are all too often omitted. Weakly-documented studies are likely suspect for use beyond the immediate research purpose, because no conclusions can be made whether the models will be able to adequately capture the impacts of substitution between inputs and substitution of technologies in electricity production.

Several works have pursued the electricity disaggregation problem without using any information regarding input costs to various generation technologies. Lindner et al. (2012) disaggregate the electricity sector in China using a random-walk algorithm with only electricity output (GWh) known. Since they do not use data on input costs, their method assumes a uniform cost of electricity to arrive at column sums, neglecting important operational realities of the sector (Hirth et al., 2014). Similarly, Shrestha and Marpaung (2006) disaggregate based solely on electricity production data.

Other electricity disaggregation efforts use available component cost data, but only leverage one aspect of the underlying economic relationships. Marriott (2007), Arora and Cai (2014), and Lindner et al. (2013) focus on row shares by allocating input costs across new technologies based largely on production (GWh) and basic assumptions (e.g. water transport is exclusive to coal-fired power and pipeline transport is split between gas and oil power). In these cases, there is no specific attention paid to the final cost structure of the technologies. These are essentially pro rata distribution-based methods. Furthermore, these are *ad hoc* methods and do not present a systematic way of introducing additional information as in the constrained optimization formulations.

Sue Wing (2008) presents a positive mathematical programming approach to incorporating cost structure and detailed engineering data (e.g. thermal efficiency, GWh production). The formulation does quite well in introducing the detailed technological data, but neglects specific attention to preserving input intensities (i.e. row shares) across the new technologies.

The discussion in Section 2.4 emphasized the necessity of preserving both cost structure and row shares in the electricity disaggregation, especially when substitution of technologies is important, as is the case for electricity generation. Given that the input cost data exist for the disaggregation task at hand, both relationships should be considered.

The example presented in this paper focuses on a disaggregation of the electricity sector in 126 regions in the Global Trade Analysis Project (GTAP) version 8 database (Aguilar et al., 2012). The United States data are presented in detail here due to the availability of reliable economic estimates in this region. The GTAP database is widely-used in the multi-region I-O, SAM, CGE, and IAM modeling communities; however, electricity is represented as a single sector which includes production, collection and distribution. Here, the original sector is disaggregated into seven new electricity sectors: nuclear, coal, gas, oil, hydro, wind, and solar power. Preserving the economic relationships in the electricity sector is necessary in this context, because such a database will be used for CGE analysis where the electricity technologies substitute as in the economic model outlined by Equations 2.3–2.8 above.

2.6.1 Economic Data for the Simple Disaggregation

The foundations of the bottom-up economic data used for disaggregation are: i) electricity production (in GWh) by technology (IEA, 2014, 2015) and ii) levelized capital, fuel, operating and maintenance (O&M) costs for each technology (IEA/NEA, 2010). Levelized costs of electricity (LCOE) are annualized unit costs (\$ per GWh). The base year for this disaggregation is 2007. These data sources are constructed from

the reported data along with some elementary assumptions. For example, missing values of LCOE are filled in the manner outlined in Peters (2015) based on regional and technological similarity.

The target \mathbf{A} matrix is constructed in a similar method as in Section 2.1 but only for a subset of the technologies for simplicity and clarity.

The balanced matrices \mathbf{X} are constructed for each of the 126 regions in the GTAP database from the initial matrices \mathbf{A} using the Kuroda, Kuroda-NC, RAS, and SPCE approaches. Only the supply-side is considered in this work. The new sub-sectors can be treated as activities in the supply of an electricity good where demanders are agnostic to the technology (e.g. perfect substitutes in demand). Alternatively, the electricity demand for each technology can be allocated based on the proportions of total cost revealed in the SPCE procedure. The results are compared using a diverse set of performance indicators from literature described in the following section.

2.6.2 Performance Indicators

It is generally useful to compare different matrix balancing methods against a variety of performance indicators, because the indicators are often highly similar to the objective function, and therefore, any individual indicator will likely be biased in favor of the associated approach (e.g. minimum percentage deviation or entropy distance). We test the four methods against four different performance indicators in the spirit of Temurshoev et al. (2011). Two of these measure percentage changes (akin to Kuroda) and two apply information metrics (akin to cross-entropy methods). Because there is no observed matrix to benchmark the balanced matrices against, each performance indicator estimates some distance between each estimated elements z_{it} of the matrix \mathbf{Z} and the *a priori* elements, z_{it}^a , of the matrix Z^a . Benchmarking with the *a priori* matrix, as opposed to some observed matrix, is an inherently weak test; however, the absence of an observed matrix is the primary motivation of the disaggregation problem. Possible alternatives are discussed in the conclusion.

(1) Mean absolute percentage error (Butterfield and Mules, 1980):

$$MAPE = \frac{1}{mn} \sum_i \sum_t \frac{|z_{it} - z_{it}^a|}{z_{it}^a} \cdot 100 \quad (2.18)$$

The scalars m and n are the dimensions of i and t , respectively. This gives an average percentage deviation between the estimated and target value.

(2) Weighted absolute percentage error (Mínguez et al., 2009):

$$WAPE = \sum_i \sum_t \left(\frac{|z_{it}^a|}{z_{\bullet\bullet}^a} \right) \cdot \frac{|z_{it} - z_{it}^a|}{z_{it}^a} \cdot 100 \quad (2.19)$$

This is similar to $MAPE$ but gives a weighted percentage deviation based on the relative size of the *a priori* element.

(3) Phi statistic (Smith and Hutchinson, 1981):

$$\phi = \sum_i \sum_t z_{it}^a \cdot \left| \ln \left(\frac{z_{it}^a}{z_{it}} \right) \right| \quad (2.20)$$

This statistic is similar to the information gain statistic given by Kullback and Leibler (1951) and entropy. For a given z_{it}^a , a z_{it} twice as large returns the same error as a z_{it} half as large.

(4) Psi statistic (Kullback, 1959; Knudsen and Fotheringham, 1986):

$$\psi = \frac{1}{z_{\bullet\bullet}^a} \sum_i \sum_t \left[|z_{it}^a| \cdot \left| \ln \left(\frac{z_{it}^a}{s_{it}} \right) \right| + |z_{it}| \cdot \left| \ln \left(\frac{z_{it}}{s_{it}} \right) \right| \right] \quad (2.21)$$

where $s_{it} = \frac{1}{2} (|z_{it}^a| + |z_{it}|)$. Knudsen and Fotheringham (1986) find both the phi and psi statistics to be some of the most accurate performance metrics.

For each of the $MAPE$, $WAPE$, and ϕ there are four different statistics: with no superscript $z_{it}^a = a_{it}$ and $z_{it} = x_{it}$; with superscript r $z_{it}^a = r_{it}^0$ and $z_{it} = r_{it}$; with superscript c $z_{it}^a = c_{it}^0$ and $z_{it} = c_{it}$; and with superscript s $z_{it}^a = v_t$ and $z_t = x_{\bullet t}$. The statistics are trivially modified for analyzing column sums so the summations do not range over the row dimension. In this way we can analyze the performance for the total matrix, row sum, cost structure, and the deviation from the implied

total costs, respectively. However, the deviations compared to the implied total costs (i.e. superscript s) should be considered less reliable in the context of this particular problem since we believe the implied total costs to be unreliable.

2.6.3 Comparison of Methods

A total of 13 statistics for each of the 126 regions provides wide coverage in evaluating the performance of the four alternative matrix balancing methods. Each region is independent of one another, so these are essentially 126 different instances of the electricity sector disaggregation. There is no clear dominance between the measures (i.e. no one performance metric always dominates another with respect to any of the statistics). Therefore, we rank the performance of each method (i.e. RAS, SPCE, Kuroda, and Kuroda-NC) for each region and average them across the 126 regions (a ranking of 1 is the best performing for that region and 4 is the worst). This ranking system provides more accurate and comprehensive information than an average of the performance indicator values themselves since the regions are truly different disaggregations. This average ranking scheme is applied to each of the 13 performance statistics. A lower score indicates better performance with respect to that particular statistic. The average rankings for each method and statistic are shown in Table 2.4 below.¹⁸

As one would expect based on the altered objective, SPCE consistently ranks better than RAS for the column and row-specific metrics (i.e. superscripts c and r , respectively). There are mixed results for the individual element error (no superscript) with SPCE performing better according to $MAPE$ and ϕ and RAS performing better for $WAPE$ and ψ .

Opposed to RAS and SPCE, Kuroda and Kuroda-NC share an objective which targets column and row shares explicitly. Regardless and still as expected, relaxing the column constraint in Kuroda-NC generally shows better performance for the column

¹⁸The complete set of 13 x 126 statistics for each method is included in the supplementary material.

Table 2.4.: Average rankings of performance indicators for each method. A ranking of 1 is best. Italics show better performing method based on objective type (entropy and SSE-type). Bold shows overall best performing. No superscript, superscript r , and superscript c indicates the metric for deviation from the individual elements, cost structure, and row shares, respectively. The superscript s is the deviation from the uncertain total cost constraint that must be imposed in RAS and Kuroda. Therefore, the metric is only used for SPCE and Kuroda-NC.

Percentage change-based performance indicators								
Method	MAPE	MAPE ^c	MAPE ^r	MAPE ^s	WAPE	WAPE ^c	WAPE ^r	WAPE ^s
RAS	3.18	3.52	2.14	-	1.31	2.88	2.27	-
SPCE	<i>3.10</i>	<i>2.97</i>	1.95	1.41	1.48	<i>2.40</i>	1.87	1.35
Kuroda	1.75	2.31	2.98	-	<i>1.85</i>	3.17	3.02	-
Kuroda-NC	1.83	1.06	<i>2.79</i>	1.56	2.96	1.40	<i>2.70</i>	1.63
Information-based performance indicators								
Method	ϕ	ϕ^c	ϕ^r	ϕ^s	ψ			
RAS	2.54	3.37	2.04	-	<i>2.13</i>			
SPCE	<i>2.33</i>	<i>2.60</i>	1.66	1.35	2.54			
Kuroda	2.73	2.72	<i>3.02</i>	-	2.02			
Kuroda-NC	2.25	1.16	3.13	1.63	3.10			

and row-specific metrics. However, Kuroda seems to perform better for all of the individual element errors with one exception. This is likely due to large swings in the column sums when the constraint is relaxed in Kuroda-NC. In fact, we see larger swings in total column sums for Kuroda-NC than with SPCE.

The large column sum swings in Kuroda-NC also greatly increases the performance in preserving cost structure; Kuroda-NC greatly outperforms all other methods. However, this seems to have some expense on the individual element error as well as the row share. SPCE performs better than Kuroda-NC for row share. It is also worth noting that Kuroda-NC outperforms SPCE with respect to individual elements for the *MAPE* metric, while SPCE outperforms Kuroda-NC when the error values are weighted by the values of the cells (i.e. the *WAPE* metric). This suggests that

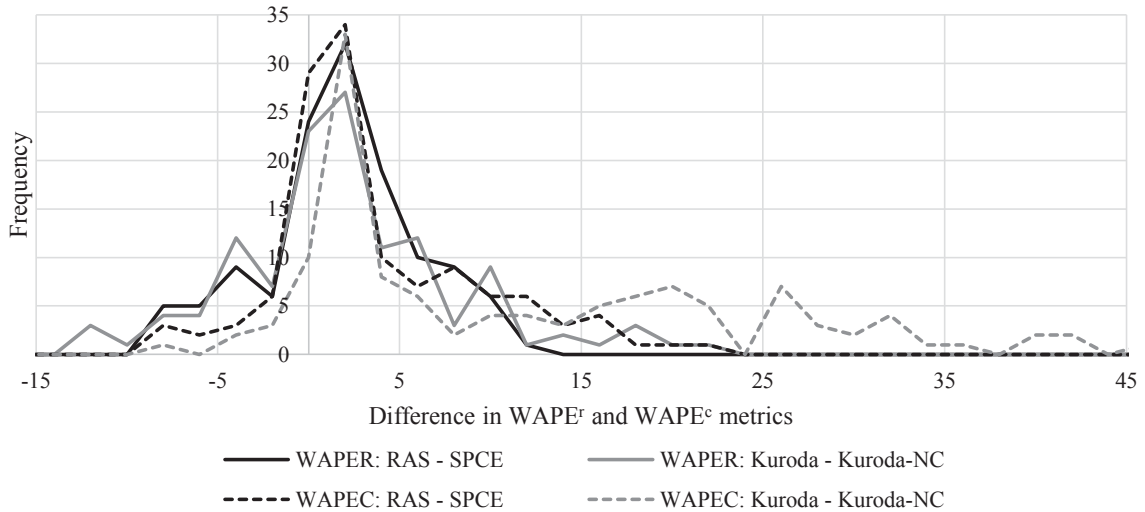


Figure 2.4.: A frequency chart of gains from relaxing total cost constraint in CE and SSE-type objectives in terms of difference in $WAPER^r$ and $WAPER^c$.

the Kuroda-NC method allows larger changes for cells with initially larger values. Temurshoev et al. (2011) reached a similar conclusion when comparing (generalized) RAS with Kuroda, which is also supported empirically here.

It is not clear if any of the four methods performs better for individual elements; there is no clear distinction. However, we can see that SPCE outperforms RAS on 8 of the 10 measures (excluding the column sum measures, superscript s , which RAS outperforms by definition). We can also observe that SPCE performs better for preserving cost structure and row share, while RAS performs better at total cost via the hard constraint. The same is true for Kuroda and Kuroda-NC. Of course, the motivation for the work is that this hard constraint is less reliable.

There is a clear trade-off between the perfect accuracy for total costs of sub-sectors using RAS and Kuroda and best preserving important economic relationships in the bottom-up data by relaxing the column constraint with either SPCE or Kuroda-NC. Figure 2.4 and 2.5 illustrate this trade-off.

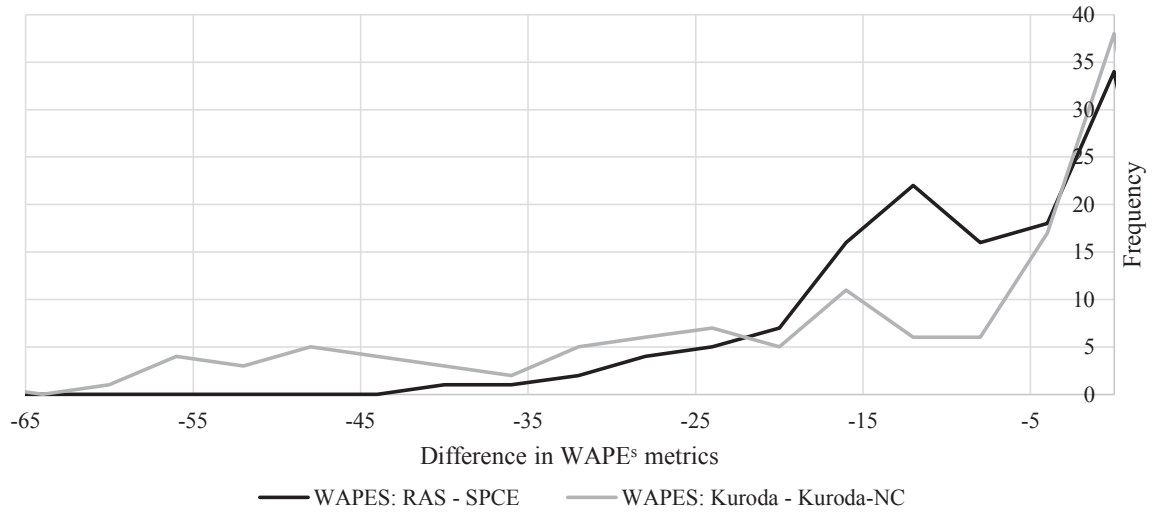


Figure 2.5.: A frequency chart of deviations from the total cost constraint when the constraint is relaxed for CE and SSE-type objectives in terms of difference in $WAPES^s$.

While there are clear gains (i.e. difference in *WAPE*-based metrics) by relaxing the total cost constraint, there are losses in terms of the column sums as defined by the constraint.¹⁹ The losses can be quite large.

However, if there is either no information on the total costs or the information is unreliable, as is the case described in this work, the losses shown in Figure 2.5 are largely meaningless. That is, how are we to evaluate error metrics if the true value is unknown and the *a priori* sums considered unreliable? The burden is on the individual researcher to evaluate which information exists, which information is more reliable, and which relationships in the economic data is most important to preserve in the balanced matrix.

It is worth noting that column constraints can still be placed on one or more of the sub-sector totals in Kuroda-NC or SPCE without compromising the methods. In this way, they are more general than their respective counterparts.

2.7 Conclusions

Matrix balancing methods are well-studied, highly effective, and efficient ways to estimate matrix elements given incomplete or even conflicting information. However, a majority of popular methods are unable to trade-off important economic relationships (i.e. industry cost structure and relative contributions to input demand) in the case where total industry costs (i.e. column sums) are unknown or unreliable. One exception is Kuroda's sum squared error-type approach, but there exists no complementary entropy-theoretic approach for this specific case. In this work we formulate a novel entropy approach to handle the case of unknown or unreliable total cost information, termed share-preserving cross-entropy (SPCE).

The SPCE formulation allows the column sum constraint to be relaxed in a constrained optimization formulation which preserves both cost structure and row shares. Previous research for incorporating additional information in the literature

¹⁹The *WAPE* is used for these figures since it seems to be the most balanced indicator between the four methods. The same insights appear using the other performance indicators as well.

are easily integrated. This formulation does not replace existing formulations. This is especially true considering the limitations of the performance evaluation using the *a priori* matrix as the benchmark for comparing against existing formulations. These formulations could potentially be evaluated against an aggregation of an observed I–O where the original observed I–O could be used as the “true” benchmark. This is outside the scope of this section because an unknown electricity-detailed matrix motivates this particular work. Instead of replacing existing formulations, this work builds on the previous work in this field and demonstrates the flexibility of the constrained optimization form of CE in situations where the data type and application preclude the more traditionally cited approaches for estimating matrix elements.

Beyond this extension, this paper implicitly advocates for the use of formulations which are designed for the specific problem at hand; the gains can be considerable, as demonstrated by the electricity disaggregation example. The formulation described in this paper could be applied seamlessly to other matrix estimation problems involving unknown column sums or where column sums have been previously constructed from simple summation. Ultimately, the preferred disaggregation method is inseparable from both the data available and the intended research objective.

CHAPTER 3. THE RELATIONSHIP BETWEEN MATRIX BALANCING AND MODELING

The previous chapter discussed an improved matrix balancing method for the case where the component costs (i.e. levelized cost of electricity) are deemed more important than the total column sums. The method relaxes the total column constraint which improves the closeness of the balanced data to the original bottom-up data.

This chapter shows that the choice of database reconciliation methodology has a significant impact on modeling results. Four commonly used disaggregation methods are compared: i) an pro rata method used by Marriott (2007), Lindner et al. (2013), and Arora and Cai (2014), ii) minimum sum of column cross-entropy (MSCCE) (Golan et al., 1994; Robinson et al., 2001), iii) RAS (e.g. Lahr and De Mesnard (2004)), and iv) SPCE described above. The experiments use identical bottom-up data to create different balanced matrices and are then taken as input to a simple partial equilibrium (PE) model which allows us to analytically trace how different disaggregation methods impact modeling results.

The modeling analysis focuses on three contemporary economic shocks. The first is a technology-specific capital subsidy (e.g. an investment tax credit). This is useful since it will highlight the value of preserving the cost structure in the sub-sectors. The second example involves a shock to the price of natural gas (e.g. a result of the US shale gas boom). Finally, a sector-wide capital tax (e.g. removal of a sector-wide tax credit) is considered. This experiment illustrates the importance of preserving row shares in the reconciled database. Model results are shown to be highly dependent on the balancing methods used to construct a CGE database and flow directly from the mathematical features of the algorithms.

In current practice, the database construction methods used in IAMs are, at best, not adequately documented. This point will only increase in importance with the increasing demand for more highly resolved analysis of critical sectors in IAMs. The results shown in this article advocate for greater introspection at the database-modeling nexus. More broadly, the results should redirect attention back to the validation of new and innovative CGE and IAM extensions. Finally, the results provide evidence that the appropriate selection of matrix balancing methods can reduce the overall deviation between bottom-up and top-down modeling.

3.1 Some Matrix Balancing Methods

The methods for disaggregation fall into two broad categories: *ad hoc* and constrained optimization. *Ad-hoc* methods employ straightforward algebraic rules to allocate the aggregate values across the different technologies. Constrained optimization methods minimize a specific distance metric with respect to important economic relationships, such as: i) cost structure (i.e. the share of an input cost in total production cost of a sub-sector, c_{it}) and ii) relative contributions to input demand from the sub-sectors, r_{it} .

Constrained optimization methods define the “closeness” metric explicitly via the objective function - seeking to find a new matrix \mathbf{X} which satisfies the column and/or row totals while coming as close as possible to the original column and/or row shares. *Ad hoc* methods target these shares implicitly. The most relevant constrained optimization methods minimize entropy distance: i) MSCCE ii) RAS, and iii) SPCE.¹ The most popular *ad hoc* method, and thus the one explored here, allocates value in the matrix based on row share alone. This is also termed a pro rata distribution method and can also be produced by removing the column constraint in the RAS formulation (Temurshoev, 2012). Other *ad hoc* methods may be equally prevalent,

¹In addition to entropy methods, Temurshoev et al. (2011) studies a number of alternate measures of deviation. For consistency and clarity in comparison we only study cross-entropy based objectives.

but are rarely publicly documented. Henceforth, *ad hoc* methods are synonymous with the pro rata distribution.

3.1.1 Pro Rata Distribution

Despite the ubiquity of well-studied matrix balancing methods, *ad hoc* approaches remain very popular in practice. The most prevalent of these is the pro rata-based allocation (Marriott, 2007; Lindner et al., 2013; Arora and Cai, 2014). Here, the matrix \mathbf{A} is comprised of the implied value from the bottom-up, engineering-level data (\mathbf{L}^0 and \mathbf{Q}^0) described in Section 2.1 and \mathbf{U} is the input employment in the original aggregate sector. The pro rata method allocates the original input value in the aggregate sector by the following equation:

$$x_{it} = \frac{a_{it}}{a_{i\bullet}} \cdot u_i \quad (3.1)$$

or equivalently, to demonstrate the uni-proportionality:

$$x_{it} = r_{it}^0 \cdot u_i \quad (3.2)$$

Of course, this is a simplification of the pro rata-based methods used by Marriott (2007), Lindner et al. (2013), and Arora and Cai (2014). Their disaggregations include more detailed inputs than the illustrative ones presented here. Basic assumptions on fuel inputs (e.g. coal to coal-fired power) and even more detailed assumptions on other inputs (e.g. water transport is exclusive to coal-fired power and pipeline transport is split between gas and oil power) are easily made. However, the general intuition is the same: row share-based allocation in the cases where no exact assumption of values on \mathbf{X} can be made. Two key points are: i) cost structure is not specifically considered in the pro rata method and ii) there can be no total cost or any other informational constraint. These can be readily implemented in the context of constrained optimization methods to follow.

3.1.2 Minimum Sum of Column Cross-Entropy (MSCCE)

The minimum sum of column cross-entropy (MSCCE), as proposed by Golan et al. (1994) and extended by Robinson et al. (2001), focuses on cost structure, but does not specifically focus on row shares. The constrained optimization problem, in its most simplified form, is as follows:

$$\min_{c_{it}} \sum_i \sum_t c_{it} \cdot \ln \frac{c_{it}}{c_{it}^0} \quad (3.3)$$

subject to:

$$\sum_t c_{it} \cdot v_t = u_i \quad (3.4)$$

$$\sum_t c_{it} = 1 \quad (3.5)$$

$$0 \leq c_{it} \leq 1 \quad (3.6)$$

where c_{it}^0 is the original cost structure implied by \mathbf{A} and where u_i and v_t are the given row and column sums, respectively, which ensure consistency with the top-down data. The optimal c_{it} result can be readily be transformed to x_{it} by multiplying them by the value of output for a given technology.

A key weakness of MSCCE is that the ordering of relative input intensities between technologies (i.e. row shares) is not always preserved (McDougall, 1999). This can have adverse consequences for economic modeling, as detailed in Section 2.4.

3.1.3 RAS

The biproportionate adjustment (RAS) method attempts to preserve both economic relationships (i.e. cost structure and row share) by targeting the elements of matrix \mathbf{A} , specifically. RAS is not always treated as a constrained optimization problem, but can be formulated with Equations 2.14, 2.11, 2.12.

This reflects a “true” cross-entropy formulation and is related to MSCCE as a weighted sum of column cross-entropy. While the divergence between \mathbf{X} and \mathbf{A}

may be greater than in MSCCE, the basic RAS solution preserves the ordering of input intensity (McDougall, 1999). Robinson et al. (2001) acknowledge that the RAS approach may be better suited to cases where both cost structure and row shares are important, as is generally the case for CGEs and IAMs.

3.1.4 Share Preserving Cross-Entropy

The MSCCE and RAS approaches both require column sum constraints (Eq. 2.11 and Eq. 2.12, respectively), whereas the pro rata approach has none. These constraints ensure a fit to an observed total cost for each sub-sector when such an observation exists. When data on the total cost associated with individual technologies does not exist, or where targeting relationships rather than totals is judged to be of more importance, the column sum constraint may unnecessarily restrict the problem (Peters and Hertel, 2015b). One convenient comparative feature for this exercise is that the result collapses to RAS when the column constraint is included. Recall, the formulation can be given by Equations 2.17 and 2.11 and explicitly balances both cost structure, c_{it} , and row share, r_{it} . This formulation is easily compared with both the pro rata and constrained optimization approaches because SPCE does not require any assumption on total cost for the sub-sectors.

3.2 Comparison of Construction Methods

In summary, there are two primary considerations when selecting a matrix balancing method: i) an objective which seeks to preserve important economic relationships (i.e. row share and/or cost structure) and ii) required constraints (i.e. total input employment (row) and/or total cost (column) constraints). Here, the required constraints are independent of additional informational constraints and refer only to requirements of the method itself. Table 3.1 shows how MSCCE, RAS, SPCE, and the pro rata approach fit into these categories.

Table 3.1.: Mathematical considerations for comparing matrix balancing methods

Consideration	MSCCE	RAS	SPCE	Pro rata
Objective	-	Row share	Row share	Row share
	Cost structure	Cost structure	Cost structure	-
Required Constraints	Total row	Total row	Total row	Total row
	Total column	Total column	-	-

Because of the interdependent relationships between the objective, the constraints, and disparities in data sources, it is difficult to reach general conclusions about an algorithm's usefulness. However, some expectations from this investigation can be formed (all of which assume no additional informational constraints and required constraints are the identical if required).

If the total cost constraint values for RAS are the same as those implied by the pro rata method, then the RAS result is equivalent to the pro rata result. However, it is worth noting that, despite this equivalence, both the RAS and SPCE allow for additional information via the constraint set. Also, as mentioned before, it is not necessary, but a total cost constraint can be imposed on SPCE. If the optional total cost constraint on SPCE is the same as the required constraint on RAS, the two methods are equivalent.

The objective function determines whether the balancing method preserves row share, cost structure, or both. The MSCCE objective considers only cost structure while sacrificing row share, and the pro rata method considers only row share, while neglecting cost structure. The RAS and SPCE objectives attempt to preserve both, but in doing so sacrifice both (although likely to a lesser degree than MSCCE and pro rata).

Required constraints may prevent an algorithm from preserving economic relationships. The row total constraint is required in all cases for CGE consistency;

however, the total cost constraint can be relaxed.² MSCCE and RAS require total cost constraints; pro rata and SPCE do not. Imposing total cost constraints may prevent the algorithm from preserving the cost structure objective because the total cost is not flexible to preserve the economic relationship of the individual elements. The row constraints impact the overall possible “closeness” of the balanced top-down data to the unbalanced bottom-up data. These constraints increasingly prevent preserving economic relationships as the constraints become increasingly restrictive (i.e. increasing disparity between bottom-up and top-down data).

Assuming no additional constraints beyond those required, the objective and constraints imply a certain ordering of how well each algorithm preserves both row share and cost structure. Here, ordering is only relevant when viewing the entirety of the matrix; ordering may not hold for individual elements. First, the pro rata methodology perfectly preserves row share while MSCCE makes no consideration whatsoever of the row shares. Therefore, SPCE and RAS lie somewhere in between. Second, SPCE will preserve cost structure better than the pro rata method, given that SPCE explicitly considers this in the objective function. Also, MSCCE should perform better than RAS with respect to cost shares, since there is no trade-off with preserving row share. These expectations are summarized later, along with the numerical results in the following sections, in Figure 3.4.

3.3 Disaggregated Matrices and Numerical Comparison

As mentioned previously, many researchers attempt to disaggregate CGE-consistent databases using detailed economic or technological data. If the bottom-up technical data and the aggregate economic data match perfectly, the balancing problem is moot; however, in practice the two data sources invariably differ, sometimes by a large margin. For example, the top-down GTAP data estimates less capital, coal, and gas employment and more O&M employment in

²The total row constraint also constrains the total value in the balanced database to the original value in the top-down data (i.e. the sum of row constraints equals the original total sector value).

the total electricity sector than the unbalanced matrix assembled directly from the bottom-up data, \mathbf{A} (Table 2.3). Therefore, the ensuing differences between the matrix balancing algorithm results can be attributed to both the balancing method (the focus of this work) and the magnitude of discrepancy between the bottom-up data and the top-down economic data.

In this section, pro rata, MSCCE, RAS, and SPCE-based disaggregations are constructed for the 129 GTAPv8 regions using the type of data outlined in Section 2.1 (i.e. annual GWh production and levelized costs fit to the GTAP input employment data). Table 3.2 shows the results for the United States for each balancing method, and Table 3.3 shows the average deviation (in absolute value) from the bottom-up data for each matrix balancing method - again for the United States.

The disaggregated electricity sectors for the United States (Table 3.2) show three main points. First, the unbalanced, bottom-up matrix, \mathbf{A} , has different total input employment values in the sector (row totals shown in Table 3.2) than the balanced matrices, all of which conform to the top-down data. However, as discussed in Section 2.3, the input employment for the balanced matrices must match that of the original electricity sector in the GTAP data. This is a major source of deviations shown in Table 3.3.

Second, expanding on the previous point, the total input employment of fuels from the bottom-up and top-down do not match. The fuel inputs are specific to a technology (i.e. coal to coal power, gas to gas power, and oil to oil power). This drives some of the deviations for the methods which attempt to preserve the cost structure of the technology because the fuel input value is inflexible.

Third, both MSCCE and RAS require a total cost constraint for each of the disaggregated technologies and are constrained to match the values implied in \mathbf{A} . However, the pro rata and SPCE methods do not require such a constraint and, in some cases, deviate greatly from the bottom-up data. This is especially true for the gas sector, a fuel-intensive technology, where the total costs of the sector are much lower for the unconstrained methods (pro rata and SPCE). SPCE has flexibility to

Table 3.2.: Disaggregated electricity sector for the United States using different approaches (A, pro rata, MSCCE, RAS, and SPCE) in 2007 US\$

A								
	Nuclear	Coal	Gas	Oil	Hydro	Wind	Solar	Total
Capital	33,985	61,938	9,744	1,313	33,091	3,103	504	143,679
O&M	16,487	27,782	5,816	2,642	4,936	855	43	58,560
Coal	0	63,625	0	0	0	0	0	63,625
Gas	0	0	84,065	0	0	0	0	84,065
Oil	0	0	0	24,823	0	0	0	24,823
Total	50,472	153,345	99,626	28,778	38,027	3,958	546	
Pro rata								
	Nuclear	Coal	Gas	Oil	Hydro	Wind	Solar	Total
Capital	28,137	51,280	8,067	1,087	27,397	2,569	417	118,955
O&M	39,869	67,185	14,066	6,390	11,936	2,067	103	141,615
Coal	0	42,782	0	0	0	0	0	42,782
Gas	0	0	47,288	0	0	0	0	47,288
Oil	0	0	0	24,111	0	0	0	24,111
Total	68,007	161,247	69,422	31,588	39,333	4,636	520	
MSCCE								
	Nuclear	Coal	Gas	Oil	Hydro	Wind	Solar	Total
Capital	32,641	32,447	8,901	1,350	39,017	3,949	649	118,955
O&M	32,388	83,151	11,147	3,746	9,978	1,150	55	141,615
Coal	0	42,782	0	0	0	0	0	42,782
Gas	0	0	47,288	0	0	0	0	47,288
Oil	0	0	0	24,111	0	0	0	24,111
Total	65,029	158,380	67,337	29,207	48,995	5,100	704	
RAS								
	Nuclear	Coal	Gas	Oil	Hydro	Wind	Solar	Total
Capital	25,991	48,392	7,039	705	33,517	2,753	558	118,955

continued on next page

Table 3.2.: *continued*

O&M	39,038	67,206	13,009	4,391	15,478	2,347	146	141,615
Coal	0	42,782	0	0	0	0	0	42,782
Gas	0	0	47,288	0	0	0	0	47,288
Oil	0	0	0	24,111	0	0	0	24,111
Total	65,029	158,380	67,337	29,207	48,995	51,00	704	
SPCE								
	Nuclear	Coal	Gas	Oil	Hydro	Wind	Solar	Total
Capital	25,991	48,392	7,039	705	33,517	2,753	558	118,955
O&M	39,038	67,206	13,009	4,391	154,78	2347	146	141,615
Coal	0	42,782	0	0	0	0	0	42,782
Gas	0	0	47,288	0	0	0	0	47,288
Oil	0	0	0	24,111	0	0	0	24,111
Total	65,029	158,380	67,337	29,207	48,995	5,100	704	

Table 3.3.: Percentage deviation (mean absolute percentage deviation across inputs and technologies) between the economic relationships before, **A**, and after balancing for the United States. Ordering in parentheses.

	Cost structure deviation	Row share deviation	Cell deviation
A	0	0	0
Pro rata	0.344 (4)	0 (1)	0.341 (3)
MSCCE	0.201 (1)	0.129 (4)	0.232 (1)
RAS	0.336 (3)	0.072 (3)	0.378 (4)
SPCE	0.315 (2)	0.044 (2)	0.326 (2)

preserve the cost structure where the value of gas input implied by the bottom-up data is higher than the gas input value in the GTAP data. Table 3.3 shows the mean absolute percentage deviations (Butterfield and Mules, 1980) from cost structure and row share for the different methods.

The ordering of mean absolute percentage deviation between the bottom-up data and the data after balancing is consistent with the expectations outlined previously. MSCCE dominates RAS and SPCE, which in turn dominate pro rata in cost structure preservation. The pro rata method perfectly preserves row share and both RAS and SPCE dominate MSCCE on this metric. Also as expected, SPCE outperforms RAS in both cases, because SPCE does not require a possibly restrictive total cost constraint. The pro rata and SPCE methods may outperform the MSCCE and RAS methods in either economic relationship if the total cost constraint is highly restrictive. As McDougall (1999) suggests, MSCCE generally preserves the original cell values better than the others.

The ordering shown for the United States in Table 3.3 generally holds for all 129 regions. Figure 3.1 shows percentage error for each region (averaged across inputs and technologies, in absolute values) between the cost structure (Figure 3.1a) and row shares (Figure 3.1b) in the balanced data and those implied by the bottom-up data for each method.

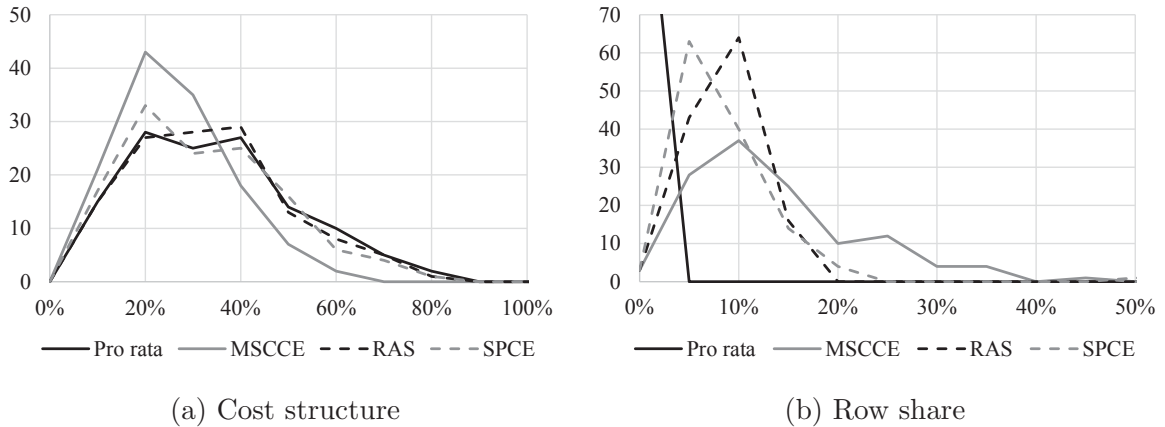


Figure 3.1.: Histograms of percentage deviation between bottom-up and balanced data in each region for both cost structure (a) and row share (b) - where deviation is the absolute percentage deviation averaged across inputs and technologies in each region.

The results across the 129 regions show numerically that the deviation between row share, r_{it} and r_{it}^0 , are generally ordered from least to greatest deviation as follows: pro rata (zero by definition), SPCE, RAS, MSCCE. The ordering for the deviation between cost structures, c_{it} and c_{it}^0 , is somewhat reversed: MSCCE, SPCE, RAS, followed by pro rata. This ordering is not necessarily identical in each region, but indicates a general tendency that is again consistent with the expectation from the mathematical structure of the matrix balancing methods.

The specific example of the United States shows what the deviation between the bottom-up data and balanced data might look like in terms of values and magnitude of deviation. The ordering of the dominance between balancing methods across the 129 regions in GTAP shows that these results are consistent with the expectations from the mathematics of the methods.

The next section demonstrates how these deviations manifest in the ensuing economic analysis based on these diverse databases. It illustrates the importance of preserving both the cost and row share economic relationships in order to ensure the model results using balanced data are as consistent as possible with the model results using bottom-up data. The shocks chosen for the simulations represent the

type of technological (e.g. shale gas extraction) and policy shocks (e.g. investment tax credits) prevalent in the electric power sector and commonly investigated using IAMs.

3.4 Economic Implications of Alternative Database Construction Methods

This section explores common energy and environmental-related shocks in the context of a model which is broadly representative of those employed in IAMs, but tractable enough to follow how economic relationships in a database map to modeling results. Therefore, a simple partial equilibrium (PE) representation of the electricity sector is presented in this section. It is for illustration only and is not an adequate representation of the electricity sector for use in a full-blown IAM, nor is it representative of the current state of electricity research.³ Rather this model is designed to clearly identify where and how differences in cost structure and row shares evidence themselves in modeling results for the electric power sector. The simple model implemented here assumes there is no trade in electricity, so the equations can be written for each region separately. Therefore, the regional index is dropped for clarity and conciseness without loss in generality. The simulations focus on the United States because of the availability of quality bottom-up data, but could be readily extended to other regions.

The model is represented in linearized form in order to highlight the role of key economic relationships, including key row and column shares. The ‘hat’ notation in Figure 3.2 refers to percentage changes in the associated levels variables. It is solved as a non-linear, initial value problem using the GEMPACK software suite (Harrison et al., 2014). The price responsiveness of electricity demand is represented via a single, aggregate demand elasticity, μ (Eq. PE1), which aggregates the demand responses of retail, commercial and industrial activities. The electricity sector production process

³There are numerous studies on the electricity sector using partial equilibrium analysis which capture a vast quantity of engineering-economic interactions. The purpose here is to precisely show the interaction between the balancing methods and model results, rather than the precise sectoral interactions.

is characterized by a quantity-preserving constant elasticity of substitution (QCES) production function which aggregates power generated from different technologies based on the CES parameter, σ , which yields a set of derived demands for electricity produced from specific technologies (Eq. PE3). Each individual power generating technology demands fuel, O&M and capital in fixed proportion to generation (i.e. Leontief production) (Eq. PE5).

Prices for electricity produced by each technology and the aggregate electricity good are assumed to cover costs, leaving no excess economic profits (Eq. PE4 and Eq. PE2, respectively) which is consistent with average cost pricing in a regulated market. Exogenous price shocks enter into the model by shifting the supply price of the input to electricity generation (Eq. PE6). The supply of inputs to the generating technologies is assumed to be perfectly elastic in this simple model (Eq. PE7).

This simple framework is used to demonstrate the effect which different supply shocks, \hat{t}_{it} , have on the model economy. Again, the bottom-up data and the PE model (Eq. PE1–PE8) are identical across the experiments. Therefore, all variation in results comes from the balancing method.

3.4.1 Simulation to Highlight the Role of Cost Structure in Modeling

Cost structure preservation comes into play when there is a shock to a particular input to a particular technology (e.g. investment tax credit for a certain technology, fuel price). In order to show the importance of preserving cost structure within each individual technology, a price shock is applied to only one sector. Table 3.4 shows the capital intensity of each generating technology in the bottom-up data (**A**) and after balancing using each method described above. The capital share in the cost structure of gas power is highlighted with a dashed box. The first simulation applies a -30% shock to the cost of capital for gas-fired power only ($\hat{t}_{i,Gas} = -30$).⁴

⁴This serves as the most straightforward example; a more relevant example in the context of IAM is presented subsequently (i.e. a negative price shock to gas as result of the shale gas extraction technology).

Table 3.4.: Capital intensity in the cost structure of technologies after matrix balancing procedures for the United States ($s_{k,t}^V$)

	$s_{k,t}^V$				
Technology	A	Pro rata	MSCCE	RAS	SPCE
Nuclear	0.673	0.414	0.502	0.400	0.416
Coal	<i>0.404</i>	0.318	<i>0.205</i>	0.306	0.316
Gas	0.098	0.116	0.132	0.105	0.096
Oil	0.046	0.034	0.046	0.024	0.034
Hydro	0.870	0.697	0.796	0.684	0.699
Wind	0.784	0.554	0.774	0.540	0.557
Solar	0.922	0.802	0.921	0.792	0.803

Production Nest	Equation	Description	No.
	$\widehat{q}e = \mu \cdot \widehat{p}e$	Final electricity demand	(PE1)
	$\widehat{p}e = \sum_t s_t^Q \cdot \widehat{p}t_t$	Final electricity price – zero profit, average cost pricing	(PE2)
	$\widehat{q}t_t = \widehat{q}e - \sigma \cdot (\widehat{p}t_t - \widehat{p}e)$	CES derived demand for individual technologies	(PE3)
	$\widehat{p}t_t = \sum_i s_{i,t}^V \cdot \widehat{p}s_{i,t}$	Price of generating technology – zero profit, average cost pricing	(PE4)
	$\widehat{q}_{i,t} = \widehat{q}t_t$	Leontief derived demand for inputs	(PE5)
	$\widehat{p}s_{i,t} = \widehat{p}_{i,t} + \widehat{t}_{i,t}$	Supply price for generating technology	(PE6)
	$\widehat{p}_{i,t} = 0$	Infinite factor supply elasticity	(PE7)
	$\widehat{q}o_i = \sum_t \widehat{q}_{i,t}$	Total input employment in electricity sector	(PE8)
<p>The hat accent designates a variable measured in percent change. $\widehat{q}e$ is the percent change in total electricity production. $\widehat{p}e$ is the percent change in price of electricity. μ is the elasticity of demand for total electricity. s_t^Q is the quantity share of production from technology t in the electricity sector. $\widehat{p}t_t$ is the percent change in price of electricity from technology t. $\widehat{q}t_t$ is the percent change in GWh production from technology t. σ is the quantity-preserving CES parameter. $s_{i,t}^V$ is the value share of input i in technology t. $\widehat{p}s_{i,t}$ reflects the supply price for input i faced by technology t. $\widehat{q}_{i,t}$ is the percent change in input i used in technology t. $\widehat{p}_{i,t}$ is the percent change in price for input i for use in technology t. $\widehat{t}_{i,t}$ is an exogenous shock to price. $\widehat{q}o_i$ is the total employment of input i in the electricity sector.</p>			

Figure 3.2.: Production structure for the simple PE model of the representative electricity sector

MSCCE is generally closer than the other methods to the capital share values implied by the bottom-up data with the exception of gas power and coal power where the deviation is comparatively large. This raises questions regarding the MSCCE method's ability to preserve cost structure despite (and probably a result of) focusing only on this in the objective.

Focusing on gas power, both the capital cost share in the RAS and SPCE approaches are closer to the bottom-up data than the pro rata approach because the pro rata method has no specific objective to preserve cost structure. The SPCE approach outperforms the RAS in this case because SPCE does not require a total cost constraint which allows additional flexibility to conform to the bottom-up data.

Table 3.5 shows that the results flow directly from the deviations from the bottom-up data. The pro rata, MSCCE, and RAS methods overestimate capital intensity ($s_{k,gas}^V$) in the gas power sector, thereby overestimating the price of gas power ($\hat{p}t_{gas}$) in Eq. PE4 and overestimating production changes ($\hat{q}t_t$) for all technologies in Eq. PE3, while the SPCE underestimates only slightly.

Table 3.5.: Targeted technology policy: a -30% shock to the price of capital for gas power in the United States

	Percent change in production (GWh) by technology ($\hat{q}t_t$)				
Technology	A	Pro rata	MSCCE	RAS	SPCE
Nuclear	-3.317	-3.957	-4.517	-3.550	-3.254
Coal	-3.317	-3.957	-4.517	-3.550	-3.254
Gas	13.017	15.531	17.730	13.934	12.770
Oil	-3.317	-3.957	-4.517	-3.550	-3.254
Hydro	-3.317	-3.957	-4.517	-3.550	-3.254
Wind	-3.317	-3.957	-4.517	-3.550	-3.254
Solar	-3.317	-3.957	-4.517	-3.505	-3.254

These results can be attributed to the deviation for this particular cell, ($s_{k,gas}^V$) of the balanced matrices. The results may not deviate for shocks to other cells; however

the implication is the same: deviations should be minimized for each cell in the matrix balancing method.

The shock to the price of capital in gas power in the United States is shown because the connection from cost structure to model results is transparent and tractable. However, the price of capital for gas is not terribly relevant in the context of energy, electricity, climate, and other policies that IAMs have proven useful for - although a similar shock on different sector might be more relevant (e.g. investment tax credit for renewables). Therefore, another relevant example is the glut of gas in the United States due to shale oil and gas extraction technology. Wellhead gas prices in the United States dropped roughly 67% from 2007 to 2012. If natural gas export terminals are constructed, then the rest of the world may also enjoy lower gas prices. A 40% decrease in gas price is applied to each of the 129 regions in the GTAP database ($\widehat{t}_{gas,gas,r} = -40$). Figure 3.3 shows a frequency chart of the absolute percentage deviation of each matrix balancing method result as compared to the bottom-up data (**A**).⁵

The MSCCE and RAS results are identical because the share of fuel in gas power is given by the employment of gas in the total electricity sector and the total cost is constrained. The total cost is flexible for the pro rata and SPCE methods, so the fuel shares may differ. The results clearly demonstrate that the SPCE, RAS, and MSCCE, which consider cost structure in their objectives, dominate the pro rata method, which does not. However, it is difficult to discern any dominance between the SPCE, RAS, and MSCCE methods. The average across technologies of the absolute percentage deviations for each balancing method are 22.83% for ad-hoc, 19.66% for RAS and MSCCE, and 18.47% for SPCE. The same results (i.e. SPCE, RAS, and MSCCE dominating ad-hoc) are found in other simulations, such as a capital subsidy for solar and wind power and a simple carbon-based tax on coal power and gas power,

⁵Due to the simple nature of the PE model, the magnitude of the price shock does not have any significant impact on the percentage deviations between the balanced and bottom-up data (e.g. Figure 3.3). That is, the frequency chart looks almost identical regardless of the magnitude of the price shock applied in each region (regions are independent from one another).

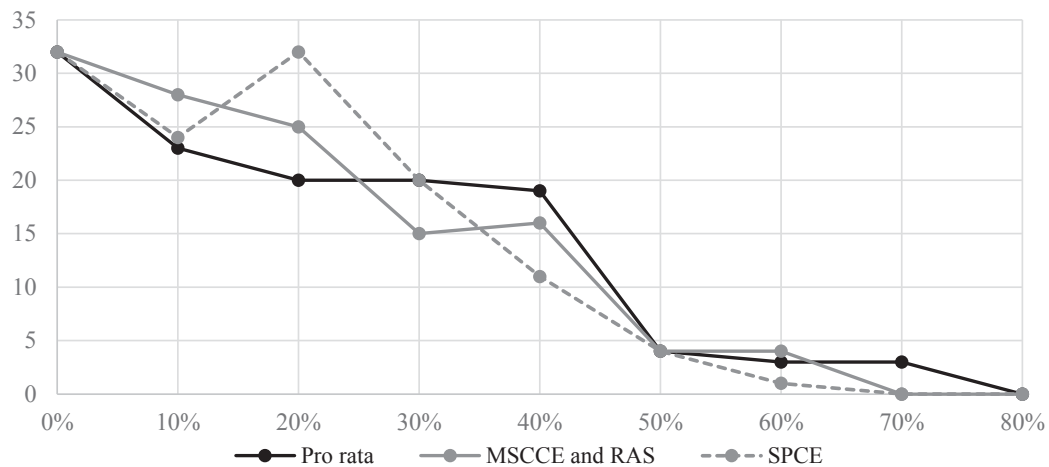


Figure 3.3.: Histogram of absolute percentage deviation from bottom-up model results and balanced data model results from a -40% shock to the price of gas in each of the 129 GTAP regions.

but to varying magnitudes. The general conclusion is that pro rata model results are less consistent with the bottom-up data model results than methods which explicitly preserve cost structure.

3.4.2 Simulation to Highlight the Role of Row Shares

Row share preservation primarily applies to a shock to an input shared by multiple technologies (e.g. investment tax credit across multiple technologies, labor taxes). Recall that Table 3.4 shows that MSCCE deviates from the bottom-up data for cost structure of capital in gas power and coal power. Another way to see this discrepancy for gas power and coal power is by capital employment across technologies (row share) in Table 3.6 below.

Table 3.6.: Capital employment across technologies (i.e. row share) after matrix balancing procedures for the United States

	Share of total capital employment in electricity sector				
Technology	A	Pro rata	MSCCE	RAS	SPCE
Nuclear	0.237	0.237	0.274	0.218	0.286
Coal	0.431	0.431	0.273	0.407	0.413
Gas	0.068	0.068	0.075	0.059	0.052
Oil	0.009	0.009	0.011	0.006	0.009
Hydro	0.230	0.230	0.328	0.282	0.215
Wind	0.022	0.022	0.033	0.023	0.023
Solar	0.004	0.004	0.005	0.005	0.003
Total	1	1	1	1	1

As expected, the pro rata method perfectly preserves the row share relationship. RAS and SPCE deviations are relatively similar which indicates the total cost constraint in RAS may not be overly restrictive in this particular case. Here, MSCCE shows a switch of ordering in row share between nuclear and coal power (shown in

bold in Table 3.6). A numerical simulation of a capital tax across all technologies is provided in Table 3.7 by implementing a uniform capital price shock of 10% (\hat{t}_{it}). A capital price shock is representative of a tax or subsidy on electricity generation investment. Investment tax credits for renewable generation is a widely used policy tool to promote renewable energy and crowd-out investment in carbon-intensive generation. Here the policy is applied to all generation types to make the connection between matrix balancing and model results clear and tractable.

Table 3.7.: Shared input policy simulation: 10% shock to the price of capital in the United States electricity sector

	Percent change in production (GWh) by technology ($\hat{q}t_t$)				
Technology	A	Pro rata	MSCCE	RAS	SPCE
Nuclear	-13.607	-5.885	-11.3	-5.779	-6.231
Coal	-0.951	-1.298	2.882	-1.266	-1.417
Gas	14.476	8.730	6.511	8.723	9.511
Oil	17.217	12.933	10.886	12.854	12.683
Hydro	-22.307	-18.800	-24.318	-18.772	-19.107
Wind	-18.557	-12.419	-23.384	-12.300	-12.755
Solar	-24.517	-23.362	-29.535	-23.464	-23.638

The positive price shock, \hat{t}_{it} , increases, $\hat{p}s_{it}$ (Eq. PE6). The ensuing impact of $\hat{p}s_{it}$ on $\hat{p}t_t$ depends on the input share s_{it}^V (Eq. PE4) which is where the differences in matrix balancing method enter the model. The matrix balancing methods affect the result of interest, $\hat{q}t_t$, via the substitution between technologies base on relative cost of technology, $\hat{p}t_t$ (Eq. PE3).

The results indicate that the large deviation in row share for coal power using the MSCCE method is translated directly to the deviation in model results. The direction of change is opposite those implied by the bottom-up data. The major implication is that, in the case of a uniform input price shock (e.g. tax break for capital

investment for renewable power) with substitutability between sectors, MSCCE can lead to opposite interpretations of model results even with the simplest of models.

3.5 Discussion

Divergent results from bottom-up and top-down modeling are well-known (Grubb et al., 1993), and there is constructive research regarding the relative merits and reducing divergence between both approaches (e.g. Böhringer (1998); McFarland et al. (2004)). The important takeaways here pertain to the reasons the results of different top-down models might diverge even in the case of identical bottom-up data. The divergence is the result of two primary factors: i) disparate bottom-up and top-down data and ii) the preservation of economic relationships (i.e. row share and cost structure) after the matrix balancing method which fits the bottom-up data to the top-down data. Of primary interest to this study are the discrepancies caused by the matrix balancing methods.

3.5.1 Disparate Bottom-Up and Top-Down Data

Moving to a CGE/IAM framework requires that the engineering data conforms to data on the circular flow of the economy, which are important in certain analyses. For example, Hazilla and Kopp (1990) and Bergman (1991) conclude general equilibrium impacts, such as input prices, output prices, and allocation of resources in the economy, can be “significant and pervasive” in the context of environmental policy. Unfortunately, the two data sources tend to differ, sometimes by large margins. The bottom-up data is constructed from levelized (i.e. annualized) costs of electricity by technology and total production while the top-down data is constructed by targeting prices of electricity, cost structure, and production data (where available) in GTAPv8. The sources and type of data are disparate.

For example, Table 2.3 shows that the share of O&M is much higher in the top-down database which draws cost away from capital and fuels. Still, in moving to a

CGE model, the balanced database must conform to the values in the top-down data via the total input employment constraint in the balancing methods described above. The constraint contributes to a large portion of the difference between the results, but is none-the-less necessary to move toward a CGE model which may be a more holistic representation of the economy as compared to the bottom-up representation.

3.5.2 Preserving Cost Structure and Row Share

Section 3.4.1 simulated a technology-specific capital price shock and a shock to the price of gas. These simulations demonstrate that preserving the cost structure for individual technologies can be important. The pro rata model does not specifically consider cost structure; inputs are allocated solely based on row share. The RAS and SPCE approaches, which specifically consider cost structure along with row share, conform closer to the bottom-up data and, therefore, the bottom-up model predictions. It is worth noting that the MSCCE may have large cost structure deviations for some technologies (e.g. coal-fired power in Table 3.4) which may be unattractive for policies targeting these technologies.

Section 3.4.2 simulated an electricity sector-wide shock to the price of capital. The MSCCE method implied an opposite result for one of the technologies. This can be attributed to the absence of consideration of the row share relationship in the MSCCE objective function. MSCCE does not specifically preserve row share, so when a shock is applied which pertains to relative input employment between sectors an opposite result may occur. Even if the result does not turn out to be opposite, it is still less convincing after observing this simulation.

3.5.3 Selecting an Appropriate Matrix Balancing Method

The decision on which matrix balancing method is most appropriate for the research task at hand depends on several factors and is highly case-specific. The initial decision is whether to include a total cost constraint. This depends on the

available bottom-up data and will drive the selection of matrix balancing method. Figure 3.4 summarizes the insights from the mathematical structure discussed in Section 3.2 which then tie this to the modeling results from Section 3.4, and charts the path to selecting an appropriate matrix balancing method for CGE and IAMs (Figure 3.5).

If there are no data on input costs to technologies, then total cost may be the only way to differentiate sectors. Also, if the researcher wishes only to shock the outputs of the new sectors (e.g. subsidy on renewable technologies) rather than the input prices in the new sectors, perhaps total cost might be preserved while sacrificing some of the component cost detail. In this case, where a total cost constraint is desired, SPCE and RAS are equivalent. A total cost constraint cannot be imposed on the pro rata method, thereby rendering it impotent for this particular information set. The numerical simulations of the RAS/SPCE and MSCCE methods show that there is no clear dominance in method in the cost structure case. That is, RAS/SPCE performs better for some sectors while MSCCE performs better for others (Table 3.4). However, RAS/SPCE performs much better in the case of the row share relevant simulations (Table 3.6) which implies that the SPCE/RAS might be the best selection when both relationships are relevant.

Alternatively, if technology specific input costs are available in the bottom-up data (e.g. levelized costs as is the case here) and the researcher wishes to shock input prices in the new sectors, the restrictive total cost constraint can be removed. In this case the applicable methods are the pro rata and the SPCE approach, because the basic MSCCE and RAS approaches require a total cost constraint. The mathematical properties imply and the numerical simulations show that SPCE performs better in terms of model results in the case of preserving cost structure (Table 3.5 and Figure 3.3) while the pro rata performs marginally better in preserving row share (Table 3.6). On balance, the SPCE seems to outperform the pro rata method when both relationships are important. It is also worth noting that pro rata methods are unable

Equivalence		
E1	RAS = pro rata	- If total cost constraint for RAS is identical to total costs implied by pro rata, then the two are equivalent. - RAS still allows for information on total cost.
E2	SPCE = RAS	- If total cost constraint is added as information to SPCE, the two are equivalent.
Cost structure preservation		
C1	SPCE > pro rata	- SPCE considers cost structure in the objective.
C2	MSCCE > RAS	- RAS sacrifices some cost structure preservation for row share. - Individual elements may differ in ordering (e.g. RAS result may be closer than the MSCCE for certain elements), but as a whole MSCCE > RAS.
C3	SPCE ~ MSCCE	- The level of restriction from the total constraint required by RAS and MSCCE will determine ordering.
Row share preservation		
R1	Pro rata > all others	- Pro rata perfectly preserves row shares.
R2	RAS and SPCE > MSCCE	- MSCCE has no consideration of row shares.

Figure 3.4.: Considerations for selecting an appropriate matrix balancing method – insights from algorithms. These only hold when no additional informational constraints are present.

Total cost constraint?	Restrictions and equivalence	Cost structure important?	Row share important?	Both relationships important?
No	- RAS not possible - MSCCE not possible	- SPCE > pro rata (C1)	- pro rata > SPCE (R1)	SPCE > all
Yes	- SPCE = RAS (E2) - pro rata not possible	- RAS/SPCE ~ MSCCE (E2, C3)	- RAS/SPCE > MSCCE (R2)	RAS/SPCE > all

Figure 3.5.: Considerations for selecting an appropriate matrix balancing method - Insights from modeling. The corresponding insight from the algorithm (Figure 3.4) are in parentheses

to leverage the vast research on incorporating additional information and reliability of information in constrained optimization (Lahr and De Mesnard, 2004).

It is more than likely that both cost shares and row shares will eventually be of importance in most CGE/IAM projects. While researchers may have a particular shock or set of shocks in mind initially, the models are often subsequently used for simulations for which it was not originally designed. Given this, the SPCE method is the most flexible method and preserves both economic relationships, thereby providing results which are the most consistent with the original bottom-up data over the largest set of shocks.

3.6 Conclusions and Broader impacts

Using a simple partial equilibrium model, the deviation between results with bottom-up data and balanced data stem from two primary sources: i) differences between the bottom-up and top-down data and ii) the matrix balancing methodology used to conform the dataset when there are disparate data. If the database implied by the bottom-up data match that of the top-down data, there is no need for the matrix balancing method at all. Unfortunately, that is rarely, if ever, the case, and the data balancing methods are necessary. This work shows that the modeling differences can be quite large based on the selection of matrix balancing method which necessitates close consideration, justification, and documentation.

This section explored four matrix balancing methods which are commonly employed to create a consistent CGE/IAM database and the implications each has on economic modeling. Their mathematical constructions (i.e. the objective and constraints) provide some insight into how they might perform in relation to two important economic relationships (i.e. cost structure and row share). The analytical investigation is supported by numerical examples in a simple disaggregation of the electricity sector. Identical data is used for each method. The numerical results are

generally consistent with their mathematical constructions regarding the economic relationships.

Numerical simulations showed the relevance of these economic relationships in modeling. The alternative balancing methods, despite identical original data, differed from the original bottom-up data results depending on their mathematical constructions and ability to preserve the economic relationships. In these experiments the original bottom-up data, partial equilibrium model, and simulations were control variables. The matrix balancing methods directly drove the modeling results.

Selecting an appropriate matrix balancing method will help decrease the divergence between bottom-up and top-down models. The SPCE method outperforms the other methods both in flexibility (i.e. it is the only method which can be used with and without a total cost constraint) and where both economic relationships are important, which is the most likely case.

The implications for large-scale CGE and IAM modeling are straightforward. First, the best way to reduce deviation introduced by the matrix balancing methods is to inform the top-down data with the bottom-up data, and vice versa. Second, in cases of disparate bottom-up and top-down data, the balancing method matters. Finally, the database construction efforts, which includes the matrix balancing, should be considered closely, justified, and documented. Moving forward data construction elements of CGE and IAM modeling efforts should be publicly documented with data and methods posted online to promote continuous improvement at the data-database-modeling nexus. This is an under-researched, but critical, aspect of IAM research and critical to the long-run credibility of this important work.

CHAPTER 4. CONSTRUCTING THE GTAP-POWER DATABASE

This work documents a tractable disaggregation methodology for the regional electricity sectors in GTAPv9 database which leverages available data and various matrix balancing techniques. The result is a transparent GTAP-Power database where specific limitations and improvements in techniques can be identified by both researchers and GTAP community members. The database will be published in hopes of continuous improvement and greater consistency in the base data amongst researchers modeling the electricity sector.

The GTAP-Power database is an extension of the GTAPv9 database in that it includes all of the data included in the GTAPv9 database. Peters (2015) provides corresponding data files and GAMS file `ely_disagg_2011.gms` which performs the GTAP-Power disaggregation for base year 2011.

4.1 Data

Recall the data used in the disaggregation for GTAP-Power are: i) electricity production (in GWh) by fuel source (IEA, 2014, 2015; EIA, 2015), ii) total value of inputs (in base year USD) to an aggregate electricity sector for each source (i.e. domestic and import), and type (i.e. basic and tax) for base years 2004, 2007, and 2011 (Aguiar et al., 2012), and iii) levelized (i.e. annualized cost per GWh) capital, operating and maintenance (O&M), fuel, and effective tax costs of electricity for select generating technologies and regions (IEA/NEA (2010); various sources). These databases are represented as matrices \mathbf{Q}^0 , \mathbf{U}^0 , and \mathbf{L}^0 with elements q_f^0 , u_{iab}^0 , and l_{ct}^0 , respectively. These data are available over an additional index, r , which covers the 140 regions in the GTAPv9 database, but this index is dropped in most of the following

notation because the regional disaggregations can be performed independently. The super-script 0 identifies these as original data sources.

The set f is the set of original technologies in the IEA database which are not differentiated based on operational characteristics (i.e. base versus peak load). The matrix, \mathbf{Q}^0 , with elements q_f^0 refers to the total electric output by each generating technology in the IEA database for each region. The EIA database was used to help fill missed regions in IEA.

The set t consists of the disaggregated sectors, transmission and distribution and all generating technologies. These are: transmission and distribution ('T&D'), seven base load technologies ('NuclearBL', 'CoalBL', 'GasBL', 'HydroBL', 'OilBL', 'WindBL', and 'OtherBL'), and four peak load technologies ('GasP', 'OilP', 'HydroP', and 'SolarP'). The matrix, \mathbf{Q}^g , with elements q_t is the expanded matrix with these new sectors for the GTAP-Power database. Electricity produced by the transmission and distribution sector is defined as the total GWh produced in the region.

The matrix \mathbf{U}^0 with elements u_{iab}^0 is an alternate representation of electricity sector in the GTAPv9 database where i is the set of all input costs to production (see Appendix A for listing), a is the set of sources (i.e. domestic or imported), and the set b is the type of cost (i.e. basic or tax).¹ The GTAP database, \mathbf{U}^0 , is used to create constraints in the GTAP-Power disaggregation.

The matrix \mathbf{L}^0 represents the levelized cost of electricity (LCOE) for each type, c (i.e. investment, O&M, fuel, own-use, and effective tax), for each new sector, t , and region.

The technologies in the IEA (\mathbf{Q}^0) and IEA/NEA (\mathbf{L}^0) do not encompass all of the technologies that are in the GTAP-Power database. The GTAP-Power database includes splits of certain generating technologies into base and peak load technologies. The intent of the split between base and peak load is two-fold. First, the total generation data (\mathbf{Q}^0) comes in the form of fuel inputs (e.g. GWh generated from natural gas); however, several different technologies (e.g. combined-cycle, combustion

¹The national version of the GTAPv9 database is created using scripts from the SplitCom application Horridge (2005). SplitCom takes the full database and creates NATIONAL and TRADE matrices.

turbine, steam turbine) are used to turn the fuels into electricity. These technologies have cost structures which must be differentiated, especially if the modeler wishes to aggregate different technologies.² Second, connecting the data to modeling, base and peak load are distinct types of generation. Without differentiating electricity production by these operational considerations, a model can have a technology like solar taking over the entire generation which is not realistic, at least in the current electricity system (i.e. without storage for time arbitrage).

The GTAPv9 electricity sector data (\mathbf{U}^0) is derived in part from the IEA GWh data. The IEA GWh data (\mathbf{Q}^0) is mapped to the GTAP regions. In the event where levelized cost data (\mathbf{L}^0) is not available for either a technology or region, averages of all available cost data are used. The accuracy of this assumption may raise eyebrows at first glance and is certainly debatable. However, considering there are only a handful of suppliers for the electricity generating units worldwide, this assumption may not be as limiting as expected in terms of both capital and O&M costs (at least for new capacity). To derive levelized costs of own-use, the value of total own-use in the electricity sector in each region comes directly from own-use in the original GTAP database, u_{iab}^0 where $i = \text{'ely'}$. The value share allocated to transmission and distribution is identical to the share allocated to transmission and distribution for the entire electricity sector (discussed later). The remainder is divided by the total GWh produced in the region to derive the electricity own-use cost per GWh. Also, estimated fuel costs, which are generally more variable by region, are derived partly from the implicit region-specific fuel prices in the GTAP database. The full levelized costs data are available in Appendix C of Peters (2015). Increasing the LCOE coverage is a major opportunity for subsequent versions.

²In the long-run specific technologies such as combined-cycle, combustion turbine, and steam turbine gas would provide a better idea of costs, but the modeling issues of how each of these technologies compete from an operational perspective is still unclear. Therefore, a simple aggregate base and peak load differentiation is a nice balance between operational considerations and data availability.

4.2 Stage 1: Base and Peak Load Split

As motivated before, this particular effort is unique to many other electricity sector disaggregations, in that the generating technologies are split into base and peak load power to reveal important cost structure and operational considerations of the electricity sector. This suggests a two-stage procedure which first identifies the GWh split between base and peak load for the generating technologies then fills the full matrix given the GWh splits. Separating the GWh split into a separate stage makes the problem more tractable and allows seamless implementation of alternative data types (e.g. detailed regional technological data) and models (e.g. Wiskich (2014)) without compromising the matrix balancing described later in Stage 2.

The base-peak load split stage minimizes the total O&M and fuel costs of base load production subject to GWh clearing constraints and an assumption that base load must account for at least 85% of total GWh produced. This is a simple way to allocate high capital, low variable cost technologies to the base load and low capital, high variable cost technologies to peak load. A straightforward improvement would be to minimize variable costs specifically; a portion of O&M costs may be fixed. The formulation is as follows and is repeated for each region, r :

$$\min_q \sum_{bl} q_{bl} \cdot (l_{O\&M,bl}^0 + l_{fuel,bl}^0) \quad (4.1)$$

subject to:

$$\sum_{bl} q_{bl} \geq \beta \cdot \sum_t q_t \quad (4.2)$$

$$q_{gasbl} + q_{gasp} = q_{gas}^0 \quad (4.3)$$

$$q_{oilbl} + q_{oilp} = q_{oil}^0 \quad (4.4)$$

$$q_{hydrobl} + q_{hydrop} = q_{hydro}^0 \quad (4.5)$$

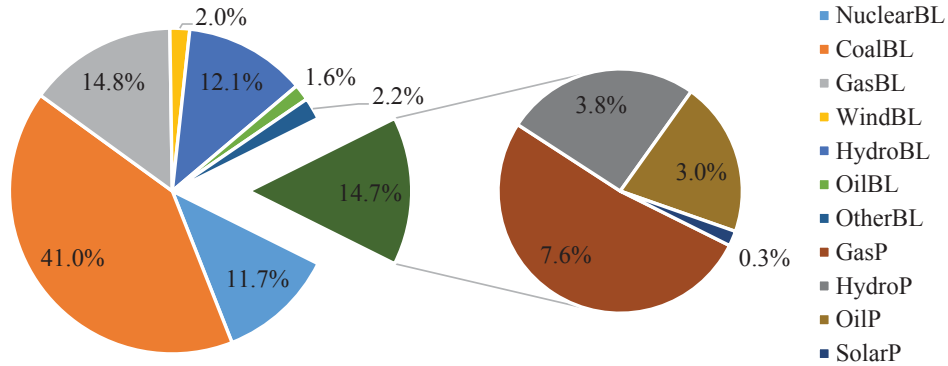


Figure 4.1.: Shares of global electricity generation from different technologies in base and peak (green cut-out) load.

where q_t is the total GWh produced by each generating technology, t . Again, q_f^0 is total GWh produced by each fuel type, f , from the IEA Energy Balance data (the dataset does not distinguish base and peak load technologies), and l_{ct}^0 is the IEA levelized cost data for each generating technology. The set t contains all generating technologies in the GTAP-Power database, and bl is the subset of t with generating technologies classified as base load power. The scalar β is the assumed proportion of base load generation in total generation (here, 85%).

One important limitation in the above method is that it cannot admit more than one technology that is both base and peak load. Alternative models which elucidate the base and peak load split (such as Wiskich (2014)) could be implemented in this stage; however, there is a trade-off between model capability, data availability, and solution improvement.

Figure 4.1 shows the shares of electricity from base and peak load technologies. Coal, nuclear, wind, and other exclusively provide base load, and solar exclusively provides peak load. The exclusive technologies have uniform levelized costs; therefore, the base and peak distinction does not have any implication on the values in the disaggregate database. Gas provides over half of the peak load. Hydro is more likely to provide base load than peak load. Conversely, oil is more likely to provide peak than base load.

4.3 Stage 2a: Targeting Levelized Cost Relationships

Peters and Hertel (2015a) show that an ideal disaggregation preserves both the cost structure and “row share” (i.e. relative input employment by technologies) implied by the economic data (in this case, \mathbf{L}^0). This is especially the case when the database will be used in a model with substitution between the electricity generation technologies. We choose to preserve these economic relationships over imposing the total cost constraints. That is, we believe the relationships to be more “trustworthy” than the total costs.

The fully disaggregated matrix is partitioned to investment, fuel, O&M, own-use, and production tax costs for transmission and distribution and each generating technology. This provides a target matrix, \mathbf{A} , based on the levelized cost and electricity production data; however, it is inconsistent with the GTAP database. Targeting relationships in levelized cost data, \mathbf{L}^0 , and fixing the other data implies that the GTAP values, \mathbf{U}^0 , as an aggregate measure, and the electricity production values, \mathbf{Q}^g , are the more trusted sources. The proposed optimization algorithm finds an estimated levelized cost which minimizes deviation from both the derived i) cost proportionality within a single generating technology (i.e. cost structure) and ii) relative input employment by generating technologies (i.e. row share) from the target levelized cost data. In doing so, the algorithm targets relationships between levelized costs rather than the levelized costs themselves. Tax costs are assumed fixed and are assigned by the value implied by the tax (\mathbf{L}^0) and production (\mathbf{Q}^g) data. The residual tax value in GTAP are allocated to the new sectors on a per GWh basis.

The objective function is designed to minimize weighted entropy distance from both the cost structure and row share relationships (Peters and Hertel, 2015b). This is termed the share-preserving cross-entropy (SPCE) method. Constraints are imposed to maintain an assumed allocation of value to transmission and distribution and ensure consistency with the GTAP database.

The target matrix \mathbf{A} , as defined in Section 2.1, is given by:

$$a_{ct} = \frac{l_{ct}^0 \cdot q_t}{\sum_c \sum_t l_{ct}^0 \cdot q_t} \cdot u_{\bullet\bullet\bullet}^0 \quad (4.6)$$

where $u_{\bullet\bullet\bullet}^0 = \sum_i \sum_a \sum_b u_{iab}^0$ or the total value in the GTAP electricity sector. The balanced matrix \mathbf{X} , as defined in Section 2.3 is given by

$$x_{ct} = \frac{l_{ct} \cdot q_t}{\sum_c \sum_t l_{ct} \cdot q_t} \cdot u_{\bullet\bullet\bullet}^0 \quad (4.7)$$

where l_{ct}^0 is replaced by the balanced levelized costs l_{ct} . The values for the linear constraints are described as

$$u_i = \sum_{i \in c} u_{i\bullet\bullet}^0 \quad (4.8)$$

Again, the index r is dropped for simplicity since each region is balanced independently. The balanced levelized costs l_{ct} are determined from the SPCE objective in Equation 2.17 above and the following constraints for each region, r :

$$\sum_t x_{ct} = u_c \quad (4.9)$$

$$\sum_d x_{dt} = \gamma \cdot u_d \quad (4.10)$$

The SPCE method is written in terms of \mathbf{X} and \mathbf{A} to conform to literature and for sake of simplicity, but are written in terms of l_{ct} and q_t in the accompanying GAMS code. The final matrix of \mathbf{L} is the estimated levelized cost which minimizes the weighted entropy distance from the economic relationships implied by the target levelized cost data (\mathbf{L}^0). The first natural log component of the objective targets cost structure, and the second targets row share. The set c consists of all the levelized costs, and subset d are the levelized costs excluding effective tax. The objective, Equation 2.17, sums across only d since the effective tax is fixed (discussed in Section 4.4.5).

The first constraint (Equation 4.9) sums over only the costs, i , which are associated with the particular levelized cost, c (e.g. labor in O&M, coal in fuel). The vector \mathbf{U}^0 is the GTAP national input value data for total value of each cost in the original electricity sector ('ely') with dimensions for source, a , (i.e. domestic or imported) and type, b (i.e. basic or tax). This ensures market clearance of the GTAP values across each levelized costs; that is, values of the new sectors in GTAP-Power can be aggregated to the GTAPv9 electricity values.

The second constraint (Equation 4.10) ensures the assumed value allocation to transmission and distribution where the scalar γ is the proportion of total non-tax value allocated to the transmission and distribution sector. The γ value does not have a great deal of literature behind it; examples of values include 4% (Marriott, 2007), 45% (Joskow, 1997), and 65% (Sue Wing, 2008) for the United States. The non-production operational expenses (i.e. transmission, distribution, customer accounts, customer service, sales, and administration) for electric utilities in the United States represent about 21% of total operational expenses (EIA (2015): Table 8.3).³ Therefore, a γ value of 21% is used for all regions in this disaggregation. In reality, the value may differ regionally which can easily be incorporated provided accurate data are available.

Additional constraints are imposed to ensure sufficient and proportional allocation of fuels into their associated technologies (e.g. total fuel costs of coal-based generation are greater or equal to the total coal costs to electricity in the GTAP database).

4.4 Stage 2b: Targeting Levelized Cost Sub-Matrices

Stage 2a returns estimated total column sums for each levelized cost (Table 6) which overcomes the unknown total costs which motivated the SPCE formulation. Therefore, RAS can be used to estimate the matrices for each levelized cost. Basic data and assumptions are used to construct the target matrices (the same \mathbf{A} as

³This does not include electricity loss in transmission and distribution. Here, we are concerned with the costs and values.

defined in Section 2.1) for each levelized cost separately. The notation is identical to the generic disaggregation problem in Section 2.3 and pertains to the relevant section only. That is, \mathbf{A} is not differentiated by any form of notation in the O&M and capital disaggregations in the following sections. They are independent disaggregations.

4.4.1 Operating and Maintenance Costs

The O&M sub-matrix disaggregation focuses on constructing a highly-detailed target matrix \mathbf{A} for the 58 costs which fall under the broad umbrella of O&M. This involves taking expert assumptions on the probability of each O&M cost, $j \in O\&M$, entering T&D or generation (GEN) (P^z) and entering a particular generation technology, t , given the cost is classified as generation (P_t^g). The O&M sub-matrix targets are defined as:

$$a_{jt} = P^z \cdot u_j \text{ for } z = \text{T\&D and } j \in i = \text{O\&M} \quad (4.11)$$

$$a_{jt} = P^z \cdot P_t^g \cdot \frac{l_{it} \cdot q_t^g}{\sum_t l_{it} \cdot q_t^g} \cdot u_j \text{ for } z = \text{GEN and } j \in i = \text{O\&M} \quad (4.12)$$

The target matrix is balanced using the RAS method. Assignment of probabilities allows the expert to integrate a great deal of specific cost-level information in a systematic manner. The final result can be seen in Table 4.4. For example, water transport is a cost only incurred by ‘CoalBL’. ‘NuclearBL’ is more skilled-labor intensive than fossil fuel technologies such as ‘CoalBL’. Also, T&D is highly service-labor intensive. The actual probabilities used in constructing this table are available in Peters (2015).

4.4.2 Fuel Costs

There are five sectors in the GTAP database which correspond to fuel costs: coal, gas pipeline, distributed gas, oil, and petroleum and coal products (‘coa’, ‘gas’, ‘gdt’, ‘oil’, and ‘p_c’ in GTAP, respectively). These are allocated using basic assumptions

and conditionals when those assumptions break down. The original GTAP coal sector is allocated to ‘CoalBL’. Both pipeline and distributed gas are allocated to ‘GasBL’ and ‘GasP’ based on the relative levelized cost between the technologies and in a manner where the proportion of types of gas are equal for each technology. The equal proportions technique is also used for oil and ‘p_c’ in ‘OilBL’ and ‘OilP’; however, petroleum-derived products do not strictly enter oil technologies (e.g. lubricants, gasoline for company vehicles). The excess ‘p_c’ is used to meet the levelized fuel cost column sum constraints for the other sectors.

Conditionals may come into play where there are fuel inputs to electricity in the original GTAP database, but there is no directly corresponding generation for a region (e.g. coal input to electricity in GTAP, but no coal generation in the OECD GWh data). The source of these residuals is case and region-specific, but may arise as a result of sectoral aggregation in GTAP, non-exclusivity of fuel use for electricity production (e.g. gas for heat in the facility), and the balancing algorithm necessary for the original GTAP database. In these cases, targets are created based on relative cost across the new sectors. High confidence in the assumptions for fuel inputs to generating technologies results in a highly constrained optimization problem.

4.4.3 Capital Costs

Although levelized capital costs only have one associated GTAP sector (i.e. capital), the difference in cost type (i.e. basic and tax) are of particular importance in the electricity sector. For instance, the US has investment tax credits which subsidize capital investments in some renewable technologies. This is an important consideration for modeling using the disaggregated database. The targets for the two type matrices are as follows:

$$a_{jt}^I = \frac{l_{it} \cdot q_t^g}{\sum_t l_{it} \cdot q_t^g} \cdot u_j \text{ for } j \in i = \text{capital} \quad (4.13)$$

$$a_{jt}^X = \frac{a_{jt}^I}{k_{jt}} \text{ for } j \in i = \text{capital} \quad (4.14)$$

where a_{it}^I is the tax-inclusive target (superscript I), and a_{it}^X is the tax-exclusive target (superscript X) for capital costs. The set j is a subset of all costs which map to capital costs ($j \in i = \text{capital}$). The scalar k_{it} is the power of the tax on capital for the electricity sub-sectors.

Entropy is minimized for a tax-inclusive and tax-exclusive matrix subject to market clearing constraints for both matrices and a total column sum for the tax-inclusive matrix. This is a similar formulation to one found in McDougall (1999).

$$\min_{x_{jt}^I, x_{jt}^X} x_{jt}^I \cdot \ln \left(\frac{x_{jt}^I}{e \cdot a_{jt}^I} \right) + x_{jt}^X \cdot \ln \left(\frac{x_{jt}^X}{e \cdot a_{jt}^X} \right) \text{ for } j \in i = \text{capital} \quad (4.15)$$

$$\sum_t x_{jt}^I = u_j \text{ for } j \in i = \text{capital} \quad (4.16)$$

$$\sum_t x_{jt}^X = u_j^X \text{ for } j \in i = \text{capital} \quad (4.17)$$

$$\sum_j x_{jt}^I = v_t \text{ for } j \in i = \text{capital} \quad (4.18)$$

where u_j^X is the row constraint for non-tax value in the GTAP data and v_t is the total column cost implied from the SPCE result for capital. The results, x_{jt}^X and $(x_{jt}^I - x_{jt}^X)$, are the balanced basic and tax matrices in the GTAP data. These are expanded to full GTAP dimensionality based on the source (i.e. domestics and imports) in the original GTAP electricity sector proportions to preserve market clearing in these dimensions.

4.4.4 Own-Use Costs

The value of total own-use in the electricity sector in each region comes directly from own-use in the original GTAP database, u_{iab}^0 where $i = \text{'ely.'}$ The total costs of own-use for the disaggregated electricity sectors is the estimated levelized cost for own-use, $l_{own-use,t}$, multiplied by the total production, q_t .

The individual electricity input costs are allocated to the new electricity sectors with the assumption that each demands identical shares of transmission and distribution and generating technologies. This is as though they draw from the grid and not necessarily the individual plant-type.

4.4.5 Effective Production Tax

The effective production tax in GTAP is labeled ‘PTAX’, which for a generating technology can be thought of as a tax on a specific type of generation, while ‘PTAX’ for transmission and distribution can be thought of a tax on electricity provision to the ultimate users. Tax costs are assumed fixed and are assigned by the value implied by the levelized tax from the data (\mathbf{L}^0) and total GWh production (\mathbf{Q}^g) data. The residual tax value (that not explained by available tax data) is additionally allocated to the new sectors on an equal per GWh basis.

4.4.6 Demand and Trade Disaggregation

The electricity mix of exports of electricity are assumed to be identical to the mix of domestic production. This assumption fills the complete trade matrix. The demand-side share allocation for each electricity sector is simply identical to the mix implied by the sum of domestic production and the net imports.

Presumably, different industries and households consume electricity from different sources depending on the sub-region and type of load. For instance, the retail industry may consume electricity predominantly during peak hours during the middle of the day. Households may consume more electricity during the peak hours immediately following school or office hours. Furthermore, households may demand renewable sources or even purchase household solar panels. Certain industries may require electricity and make long-term agreements for base load electricity.

Unfortunately, anything beyond pure assumption is currently unavailable for this research. The disaggregation of the demand-side in this work assumes all

users demand identical shares of transmission and distribution and each generating technologies.

Alternatively, the transmission and distribution sector could be separated from generation where generation would be sold to transmission and distribution, and users would purchase the transmission and distribution. This may make some sense in terms of how the electricity sector operates (at least in the United States); however, this type of database construction does not allow for different generation demands by industry. The database construction described above is general enough to allow for this; although due to data limitations, uniform mixes across industries are assumed for this particular version.

4.5 Results

The final results of the supply-side disaggregation are presented in this section. The demand-side is less interesting because of the lack of available data. The data and assumptions explained above are available upon request. This section focuses on the error between the estimated levelized costs and the IEA/NEA data and how important features in the original datasets are captured in the disaggregated data.

4.5.1 Pro Rata Method for Large Deviations

For some regions the deviation between the estimated and the target levelized costs can be quite large. While deviation is expected, large deviations may indicate broader issues in the OECD electricity production and, more likely, the electricity sector in the GTAP database. For instance, the target estimate of capital requirements for electricity in Cambodia derived from the levelized costs and production data is 75.9 million USD; however, the GTAP database reports only 1.8 million USD of capital is allocated to the entire electricity sector. Obviously, there is some discrepancy in reporting between the top-down and bottom-up type datasets.

To accommodate for largely disparate data, an additional bound constraint is added to the Stage 2a formulation which bounded the estimated capital and O&M levelized costs for each generation technology to n and $(1/n)$ times the target levelized cost value threshold. Fuel costs are excluded from these bounds because these are generally directly mapped to fuel input costs in GTAP (e.g. coal to ‘CoalBL’) and the costs are highly variable across regions which limits the relevance of using an average of available levelized cost data as an initial estimate. A single capital and O&M levelized cost to a generation technology which deviates by n or $(1/n)$ times the target estimate results in an unsuccessful completion. The number of successful completions in total share of GWh terms are shown in Figure 4.2. Beyond a certain threshold (x-axis) it may be better to allocate each levelized cost using an *ad hoc* method. The threshold chosen is 10 because at this point over 95% of the global GWh produced converges using the SPCE method.⁴

The pro rata method is used for those regions that do not converge by allocating each costs by the production weighted levelized costs described by for each unsuccessful region recalling Equation 3.1.

This does not specifically preserve the cost structure, but the data is so disparate in these regions the any modeling of these regions individually is suspect to begin with.

The jump in GWh converging from a bound of 9 to 10 in Figure 4.2 is due to the convergence of Russia at a bound of 10. The GTAP value of capital in Russia is much lower than the value implied by the target levelized cost of capital. Later, this section discusses how the GTAP database construction may be able to leverage the levelized cost data to eliminate such large discrepancies moving forward while recognizing the limitations in the target cost data as well.

⁴With a threshold of a 10 or 1/10 times deviation from the original levelized cost value the following 22 regions cannot be reconciled: Oman, Rest of Oceania, Brunei Darussalam, Laos, Rest of South Asia, Argentina, Ecuador, Honduras, Estonia, Lithuania, Belarus, Rest of Former Soviet Union, Georgia, Bahrain, United Arab Emirates, Rest of West Asia, Guinea, Rest of West Africa, Kenya, Madagascar, Malawi, and Rest of Southern Africa. These regions produce less than 3% of the world’s electricity and in most cases would be aggregated into larger regions.

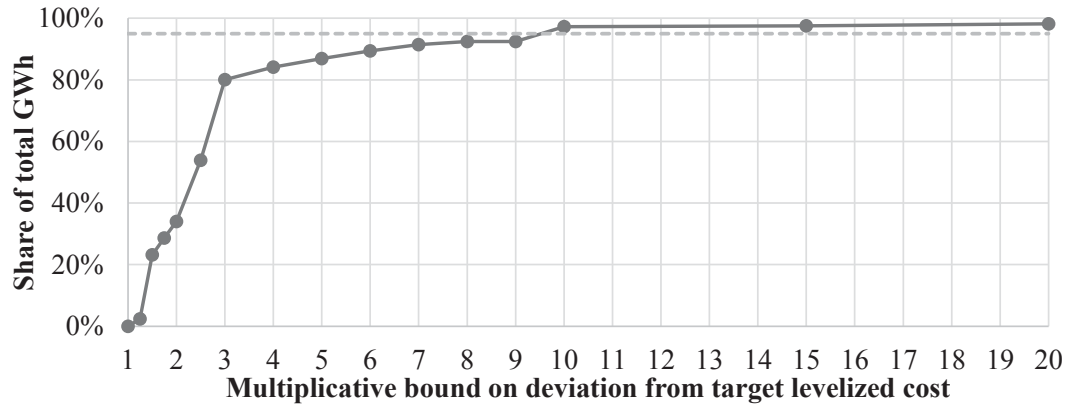


Figure 4.2.: Share of total GWh which converge using the SPCE procedure with the bound of deviation from target LCOE data

4.5.2 Deviation from Target Levelized Cost Data

It is worth reiterating that the procedure described above implies that the GTAP values, as an aggregate measure, and the electricity production values are a more trustworthy source than levelized cost, as a stylized representative of actual costs determined from a number of assumptions. This is why we fix the GTAP, \mathbf{U}^0 , and the estimated GWh production, \mathbf{Q}^g , values and target the levelized costs, \mathbf{L}^0 .

Table 4.1.: Percent deviation from non-fixed target levelized cost for each generating technology for the United States. The average absolute deviation for non-fixed values is 21.3%.

LCOE	T&D	Nuclear	Coal	GasBL	Wind	Hydro	Other	GasP	Oil	Solar
Capital	-10.5%	5.0%	10.0%	-16.9%	-4.1%	-7.5%	-5.4%	-17.2%	-16.1%	-9.6%
Fuel	-19.2%	-	63.0%	-25.1%	-	-	-14.6%	-25.1%	-24.2%	-
Own-use	-8.7%	7.0%	12.0%	-15.2%	-2.1%	-5.6%	-3.5%	-15.5%	-14.4%	-7.7%
O&M	34.0%	57.0%	64.0%	24.0%	43.0%	38.0%	41.0%	41.0%	24.0%	35.0%
PTAX	-	-	-	-	-	-	-	-	-	-

Table 4.2.: Percent deviation from non-fixed target shares of levelized cost in the total cost (i.e. cost structure) of each specific generating technology in the United States. The average absolute deviation for non-fixed values is 16.7%.

LCOE	T&D	Nuclear	Coal	GasBL	Wind	Hydro	Other	GasP	Oil	Solar
Capital	-21.9%	-19.1%	-19.2%	3.0%	-9.3%	-5.9%	-7.9%	4.0%	3.0%	-3.6%
Fuel	-29.5%	-	20.0%	-7.1%	-	-	-16.8%	-6.1%	-6.6%	-
Own-use	-20.3%	-17.5%	-17.5%	5.0%	-7.4%	-3.9%	-6.0%	6.0%	6.0%	-1.7%
O&M	17.0%	21.0%	21.0%	54.0%	35.0%	40.0%	37.0%	37.0%	55.0%	44.0%
PTAX	-12.7%	-23.1%	-26.6%	24.0%	-5.4%	2.0%	-2.6%	2.6%	25.0%	7.0%

Table 4.1 shows the percentage deviation of the estimated levelized costs from the target levelized costs in the United States. The average deviation for non-fixed levelized costs is 20.9%. The estimated levelized cost for ‘NuclearBL’ and ‘CoalBL’ are larger than the target levelized cost while the majority of others are lower. It is also evident that the O&M cost in the GTAP data is much larger than the cost implied by O&M in the levelized cost; the deviation is highest for O&M costs and the balanced estimates are all larger than the target levelized costs. The histogram in Figure 4.3 shows how these deviations are distributed for all regions, and Figure 4.4 shows the different between OECD and non-OECD countries.

Table 4.2 and Table 4.3 show the deviation from the cost structure and relative cost intensity for the United States, respectively. Again, the disparity between the target levelized costs and the GTAP data in O&M is the primary source of deviation in cost structure. However, the relative fuel costs seem to be the primary source of deviation in the row share. The effective tax has no deviation because the taxes are allocated on a row share basis. The average deviation is 16.7% and 12.3% for cost structure and row share, respectively.

Deviations can be attributed to two primary factors: i) discrepancies in the values implied by the different data sets and ii) assumptions made in the procedure itself. An example is the O&M levelized costs for the United States (discernable in all three deviation tables). While the error in fuel and tax estimates are relatively low, the estimated levelized costs for O&M are much higher than the values in the IEA/NEA dataset (Figure 4.3). This indicates that the O&M costs implied by original GTAP dataset are much higher than the IEA/NEA data. However, this deviation can also be attributed to our assumption of the cost structure of the transmission and distribution sector. If this assumption is altered to include a larger share of O&M in the total cost of transmission and distribution, some of the ‘excess’ O&M in generation will be absorbed by the sector. The own-use cost has a high deviation because the target for transmission and distribution was constructed from the ‘similar-to-communications-sector’ assumption and targets for generating

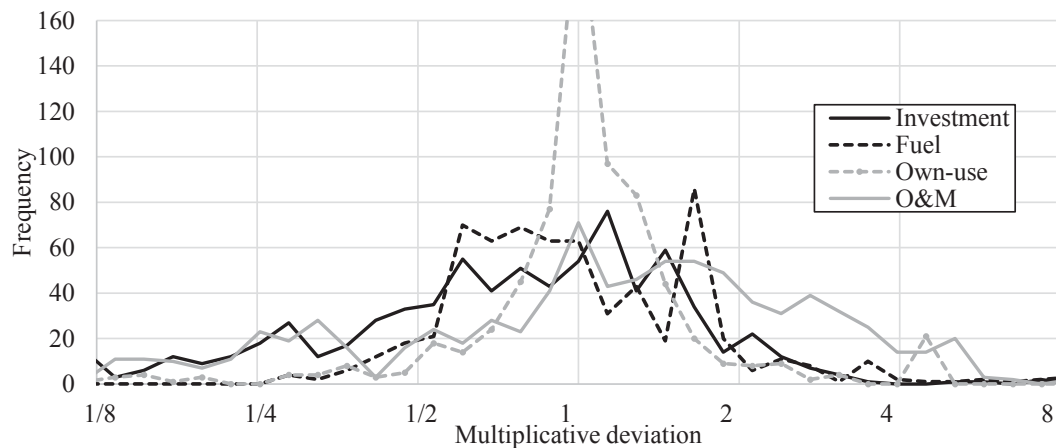


Figure 4.3.: Histogram of deviation between estimated levelized cost and target levelized costs for all regions. For each distribution, the deviation of the median from one indicates bias and larger standard deviation indicates larger deviation between the disaggregate data and original GTAP data.

Note: l_{ct}/l_{ct}^0 plotted on a log-scale.

technologies were constructed from GWh data, and the targets do not necessarily sum to the total own-use for electricity in the GTAP data. While the error values for the United States and some other OECD countries are relatively low, the errors can be quite high for regions where the GTAP and IEA electricity production data are questionable and where the levelized cost is derived through averages.

Figure 4.3 shows that O&M costs are skewed to the right which indicates that GTAP, in general, has more O&M cost than the levelized cost data. However, at the left-hand extremum for both investment and O&M costs, Figure 4.3 shows that for some regions and technologies there is significantly less value in the GTAP data than what is implied by the levelized cost data. In other words, the estimated levelized costs are lower than the target data set. This could be partly a result of averages from mainly OECD countries used as levelized cost in developing and other low-income countries where no data are available (see Figure 4.4). For instance, a low-income country may face significantly less labor costs, which is a major component of O&M costs. Another disparity could be between the assumptions of parameters used to construct the levelized costs and similar assumptions in the GTAP database. The nature of the deviations, shown in Figure 4.3 and Figure 4.4, implies that the levelized cost data can be improved greatly. Furthermore, high-fidelity cost data could also lead to improvements in the construction of the GTAP electricity sector itself.

4.5.3 Main Result

All of the original 62 costs were disaggregated using the method described above, and the results are then aggregated to 21 sectors for analysis (See Appendix A in Peters (2015) for sectoral mapping). Table 4.4 shows the input values to the disaggregated sectors for the United States. The values are the sum of sources and type dimensions. With the exception of capital subsidies to solar power ('SolarP') and balancing of the capital across the other users, the hidden dimensions are allocated in identical proportions.

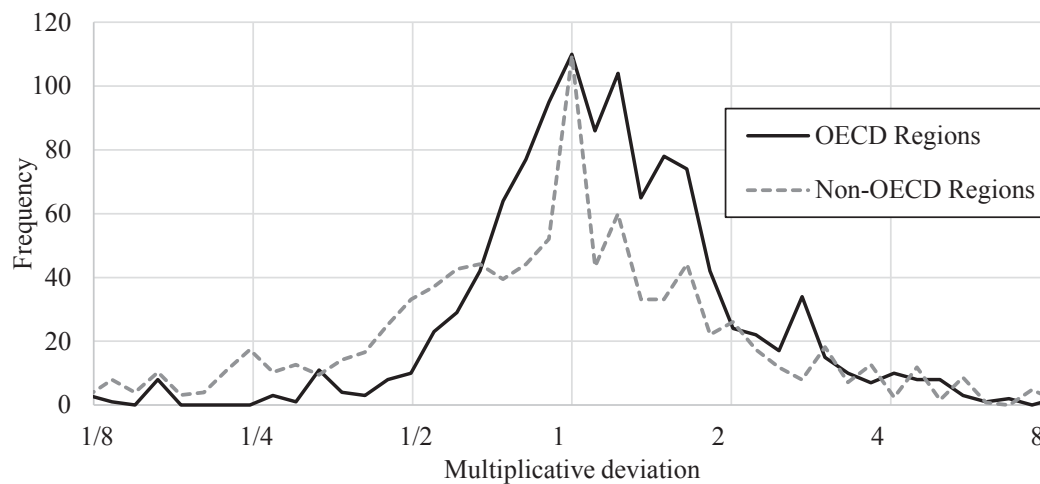


Figure 4.4.: A histogram comparing deviation between estimated levelized costs and target levelized costs for OECD countries and non-OECD countries. The larger mass of OECD regions around one indicates a closer match between disaggregate data and original GTAP data.

Note: l_{ct}/l_{ct}^0 plotted on log-scale. Non-OECD counts (1,408 non-fixed values) are scaled to OECD counts (1,111 non-fixed values) for comparison.

The fuel sectors (i.e. coal, gas, gas distribution, oil, and petroleum products) are allocated to the corresponding generating technology. Coal enters ‘CoalBL’, Oil enters ‘OilP’ only as there is no GWh generated from oil technology as base load in the United States. Gas and gas distribution enter in equal proportion to ‘GasBL’ and ‘GasP’. However, the proportion of gas fuels in ‘GasP’ to gas fuels in ‘GasBL’ is greater than the proportion of GWh in ‘GasP’ to ‘GasBL’ due to a higher levelized cost of fuel for peak gas production. The opposite is true when looking at capital to the gas generating technologies because ‘GasBL’ is more capital intensive than ‘GasP’. A portion of petroleum products also enter ‘GasBL’ and ‘GasP’ in order to reach the levelized cost target (i.e. the total gas inputs in GTAP were insufficient). The petroleum and coal products sector in GTAP contains many different energy fuels (e.g. coke, refinery gas, diesel), so it is difficult to distinguish the actual composition of this sector. As discussed previously, some of these energy fuels may very well enter alternative types of production other than strictly oil technologies. These also enter in fixed proportion between gas technologies. Similarly, the relative levelized cost intensities between technologies is preserved when we look at the other levelized costs and generating technologies as well.

The probability tables, P_t and P_g , used in the disaggregation can be found in Appendix B. Focusing on two O&M sectors which had no additional assumptions beyond relative cost intensity between technologies, chemicals & rubber and non-ferrous metals, we see that the relative costs are similar across the technologies. The ratio of value of chemicals and rubber to non-ferrous metals is approximately 10.7 for each technology.

However, general assumptions can be made about many O&M sectors. First, water transport is allocated strictly to ‘CoalBL’ (i.e. $P_t(GEN) = 1$ and $P_g(CoalBL) = 1$), since coal is generally the only fuel source which is transported domestically by waterway in the United States. Second, a $2/3$ probability of $P_t(T\&D)$ was made for various sectors in the services set under the assumption that a majority of the sales, customer service, etc. of the utilizes fall under these sectors in transmission

and distribution. This is a simple and somewhat arbitrary value, but demonstrates the ability to add expert intuition into the methodology. A similar method can be adopted to redistribute skilled and unskilled labor. This may require a balancing act between relative probabilities between types of labor within a technology and across technologies. The complex allocations of these two labor types in generation demonstrate how some of these assumptions may sacrifice transparency of the final results. These results show that the skilled to unskilled labor ratio is higher for 'NuclearBL' and renewable sources (i.e. 'WindBL' and 'SolarP') than fossil-fuel based generating technologies.

Table 4.4.: Example result in the GTAP-Power database: an electricity-detailed disaggregation of the GTAP electricity sector for the United States in 2011. A 24 sector consolidation of the full 63 sectors. Source: Peters (2015)

	TnD	Nuclear	Coal	GasBL	Wind	HydroBL	OilBL	Other	GasP	HydroP	OiIP	Solar	Total
Total electricity production by generating technology in GWh													
Total production		821,405	1,872,215	445,135	120,854	321,733	0	96,289	611,425	0	31,416	6,153	4,326,625
Share		19.0%	43.3%	10.3%	2.8%	7.4%	0.0%	2.2%	14.1%	0.0%	0.7%	0.1%	100%
Cost of inputs to electricity sectors in millions of 2011 US\$													
1 Coal	-	-	61,781.0	-	-	-	-	-	-	-	-	-	61,781.0
2 Gas	-	-	-	2,131.0	-	-	-	-	3,953.0	-	-	-	6,084.0
3 Gas distribution	-	-	-	9,248.0	-	-	-	-	17,150.0	-	-	-	26,398.0
4 Oil	-	-	-	-	-	-	-	-	-	-	13.0	-	13.0
5 Oil & coal products	12.9	-	-	5,563.0	-	-	-	2,651.0	10,316.0	-	5,081.0	-	23,623.9
6 Agriculture & food	4.0	2.8	2.8	0.2	0.3	0.6	-	0.3	0.4	-	0.1	0.0	11.4
7 Mining	1.1	1.7	0.2	0.0	0.0	0.0	-	0.0	0.0	-	0.0	0.0	3.0
8 Non-metal minerals	13.2	9.1	9.2	0.7	1.0	2.0	-	0.9	1.2	-	0.3	0.0	37.5
9 Paper & textiles	170.0	117.1	118.9	8.9	12.4	25.6	-	11.9	14.9	-	3.8	0.6	484.1
10 Chem. & rubber	520.0	358.0	364.0	27.2	37.9	78.4	-	36.3	45.7	-	11.7	1.9	1,481.1
11 Non-ferrous metals	95.5	65.8	66.8	5.0	7.0	14.4	-	6.7	8.4	-	2.2	0.3	272.0
12 Ferrous metals	102.1	70.3	71.4	5.3	7.4	15.4	-	7.1	9.0	-	2.3	0.4	290.8
13 Water transport	-	-	1,371.0	-	-	-	-	-	-	-	-	-	1,371.0
14 Air & land trans.	4,062.0	3,380.0	1,803.0	257.2	357.4	739.0	-	342.6	431.1	-	110.7	17.9	11,500.9
15 Retail	1,319.0	908.0	922.0	69.1	96.1	199.0	-	92.2	116.0	-	29.7	4.8	3,755.9
16 Services	25,801.0	13,080.5	13,280.0	994.7	1,383.5	2,862.1	-	1,327.0	1,669.5	-	428.1	69.4	60,895.8
17 Cons. & machinery	1,787.6	1,231.0	1,249.5	93.6	130.6	269.6	-	124.8	156.7	-	40.3	6.5	5,090.4
18 Tech-aspros	456.0	631.0	320.0	24.0	66.7	69.0	-	32.0	80.5	-	20.7	1.7	1,701.6
19 Clerks	4,097.0	705.0	716.0	53.6	74.6	154.0	-	71.6	90.0	-	23.1	3.7	5,988.6
20 Service-shop	204.0	134.0	273.0	20.5	14.2	58.8	-	27.3	34.3	-	8.8	0.7	775.6
21 Off_mgr-pros	11,862.0	3,425.0	1,159.0	86.8	121.0	250.0	-	116.0	146.0	-	37.4	6.1	17,209.3
22 Ag_othlowsk	7,822.0	5,386.0	5,468.0	410.0	570.0	1178.0	-	546.0	687.0	-	176.0	28.5	22,271.5
23 Capital	22,159.0	23,626.0	40,623.0	3,409.0	7,004.0	26,577.0	-	2,140.0	3,002.0	-	298.0	1,129.0	129,967.0
24 PTAX	5,210.0	-4,753.0	8,481.0	2,016.0	-107.0	545.0	-	72.0	2,770.0	-	142.0	-7.0	14,369.0
Total input cost	85698.5	48,378.3	138,079.7	24,423.9	9,777.1	33,037.9	-	7,605.7	40,681.6	-	6,429.1	1,264.6	395,376.4
Share	21.68%	12.24%	34.92%	6.18%	2.47%	8.36%	-	1.92%	10.29%	-	1.63%	0.32%	100%

Note: Rows 1-5 are associated with fuel costs, 6-22 with O&M costs, 23 with capital, and 24 with effective tax. Own-use is not shown here. May not sum to totals due to rounding.

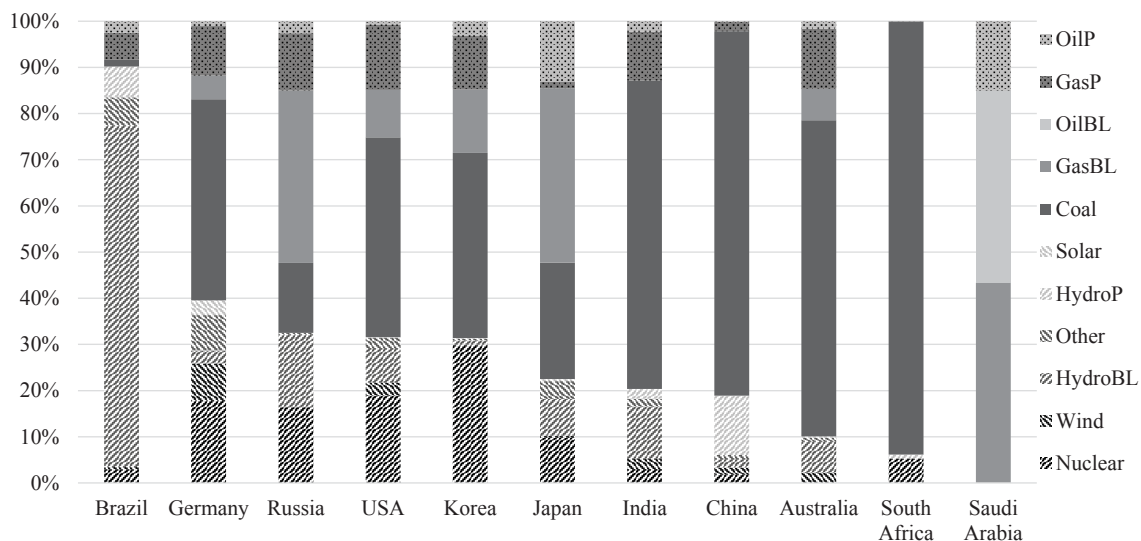


Figure 4.5.: The share of GWh produced by each technology in GTAP-Power for the same ten largest fuel-based electricity sectors plus Brazil as Figure 4.6 (ordered left to right by share of non-fuel based technology share). The non-fuel based technologies represented with diagonal lines (i.e. Nuclear, Wind, HydroBL, Other, HydroP, and Solar) are only implicitly represented by ‘Capital’ in the original GTAP database.

Moving from the US electricity to the global level, Figure 4.5 shows that non-fuel based technologies play a large role in the electricity production for many countries (i.e. nuclear, wind, hydroelectric, solar, and other). In the original GTAP database, these technologies would be agglomerated with the rest of the capital in the electricity sector. With the electricity-detail in GTAP-Power, it is possible to distinguish these very different technologies in a CGE database and provide a better representation of the sector to model of electricity, energy, and climate policies using social accounting and CGE methods.

Looking just at the fuel-based technologies in Figure 4.6, we see the share of fuel-based technologies as well as the import/domestic share varies greatly between countries. Further, Figure 4.6 shows that almost all of Korea’s coal and gas power uses imported fuels. The source of the imports are countries that use a significant amount of fuels in their own electricity sector (e.g. Russia, US, China, Australia). These

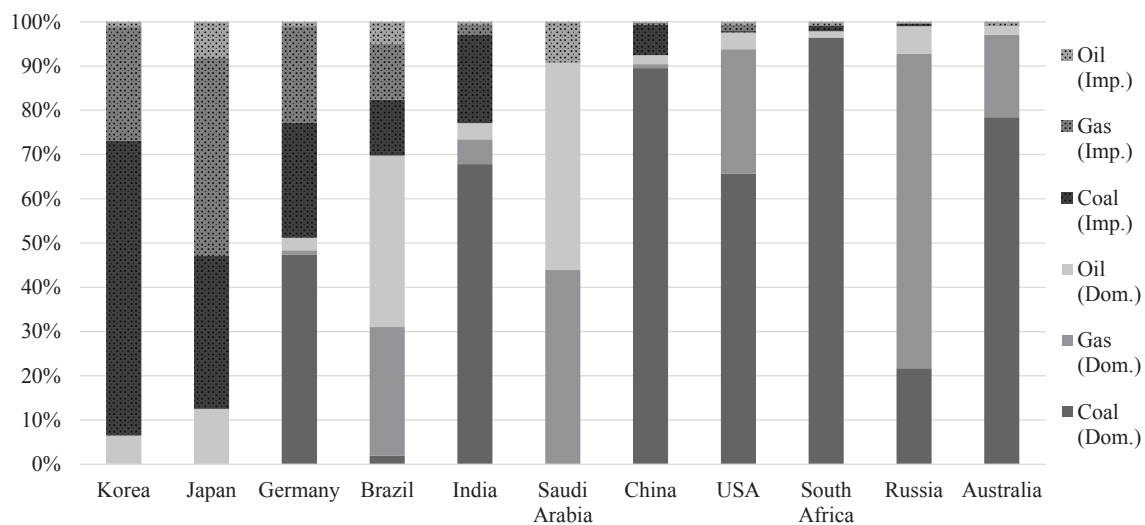


Figure 4.6.: The share of domestic and imported fuels used in the electricity sector for the ten largest fuel-based electricity sectors (ordered left to right by import intensity) plus Brazil. Both the fuel and source composition differ greatly between countries.

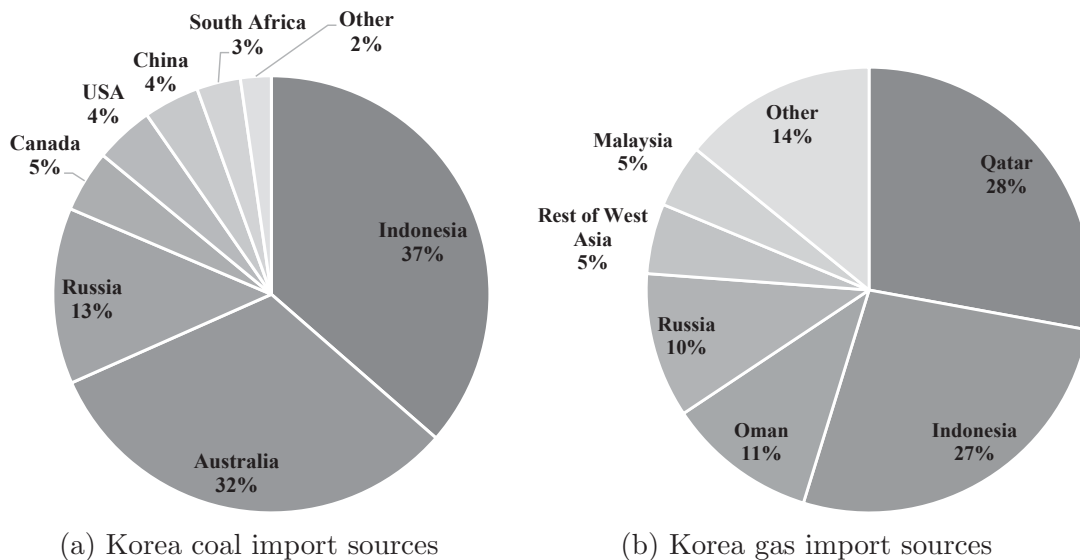


Figure 4.7.: Korea produces more than 32% of its electricity from imported fuels. These charts show the composition of Korea's coal and gas imports.

two charts show that energy trade is partly driven by domestic electricity sectors. GTAP-Power allows energy trade to be understood in greater detail.

4.6 A Look Back at the GTAP Database Construction

There may exist some opportunity to reconcile the aggregate electricity value implied by the disaggregated levelized cost data with those from original aggregate GTAP electricity sector. Table 4.5 below shows the aggregate value of inputs to the electricity sector in the United States implied by the disaggregated data compared to those in the GTAP aggregate electricity database. The latter is a constraint on disaggregation, so it is also identical to the aggregate electricity sector in the GTAP-Power database. This section presents *ideas*, as opposed to guidelines, on how such a reconciliation might be performed in subsequent versions.

On one hand there is the bottom-up data constructed from levelized costs and production levels. Recall that due to the heterogeneous reality of electricity markets these levelized costs can be misleading in many ways (Joskow, 2011; Hirth et al., 2014).

Table 4.5.: Deviation between total aggregate inputs to the electricity sector implied by the disaggregate data (used as targets) and the original values in the GTAP ‘ely’ sector (used as consistency constraints) for the United States.

Total aggregate inputs to electricity (millions of 2011 US\$)			
LCOE	Disaggregate data	GTAP ‘ely’ sector (u_c)	Deviation
Inv	131,763	129,967	1.4%
Fuel	112,309	117,899	-4.7%
Own-use	26,002	26,002	0.0%
O&M	92,575	133,139	-30.5%
Tax	14,370	14,370	0.0%
Total	377,020	421,378	-10.5%

On the other hand the aggregate GTAP electricity sector is constructed from targets that are derived from various sources, namely contributed I–O tables and IEA energy data. Even in the long-run, it is unlikely that the I–O tables contributed by the GTAP community will include all or even some of the electricity sub-sectors described here. Many contributions may not even include a separated electricity sector. Therefore, the new electricity sectors described here will likely remain a disaggregation of an aggregate electricity sector in the main GTAP database construction.

Therefore, despite the known limitations, there may be opportunity to use the levelized cost data to create targets for the aggregate electricity sector in GTAP, especially where quality data may not exist to target the sector otherwise. This would help reconcile the bottom-up and top-down perspectives of the electricity sector.

There are at least three distinct cases in a possible aggregate electricity reconciliation exercise. For each GTAP region: i) the levelized cost data are more “trustworthy” than the GTAP target, ii) the GTAP target is more “trustworthy” than the levelized cost data, or iii) they are equally “trustworthy” (or equally “untrustworthy”). In the first two cases, it may be best to simply use the target

the researcher deems more “trustworthy”. However, there should be some additional introspection when these deviate by such large margins.

The third case may be more interesting. Typically, the quality of the data follows the collection efforts of the region and both the bottom-up and top-down data are either “trustworthy” or “untrustworthy,” rather than the two cases described before. Still differences inevitably arise, as shown for the United States in Table 13. In this case, there might be two options based on the cost structure component of the data. If the cost structure of the aggregate GTAP electricity sector is “trustworthy”, a simple average of the two data sources for each input cost to the aggregate electricity sector could suffice. What might be a more likely case, is that the top-down total value in the electricity might be accurate since it can be easily constructed from a price of electricity and total production (demand-side), but the cost structure is created from assumption rather than data. Here, the targets for inputs to the aggregate electricity can be constructed by taking the total electricity sector value from the top-down data and imposing the *aggregate* cost structure (GWh-weighted average levelized cost plus T&D) implied by the bottom-up data.

These methods might help decrease the gap between the bottom-up and top-down modeling using the GTAP-Power database. This section documents ideas gathered from this particular disaggregation exercise and not necessarily the path GTAP will continue in the future.

4.7 Conclusions and Path Forward

Chapters 2–4 discussed advances in the construction of global economic databases. A novel matrix balancing method that preserves important economic relationships is developed in Chapter 2. The new method outperforms commonly-used methods when the bottom-up cost data are more reliable than the total sub-sector cost data. Chapter 3 shows that this can have important consequences in modeling using a balanced

database. These are used in this chapter to develop the GTAP-Power database, the foremost publicly-available electricity-detailed general equilibrium database.

GTAP-Power is a disaggregation of the GTAP electricity sector into transmission and distribution, base load generating technologies, and peak generating technologies for use in CGE models. The method leverages available data and reasonable assumptions to construct the database in a replicable and transparent manner. Application to CGE and integrated assessment models which are built on the GTAP database is straightforward.

The resulting electricity-detailed GTAP-Power database can be used by researchers to advance modeling of electricity, energy, and climate policies using social accounting and CGE methods.

An additional motivation for this work is to identify strengths and limitations in database construction for consistency and continuous improvement in the GTAP community. There are many limitations to this work that offer opportunity for continuous improvement given additional data sources. Some of these are listed below:

1. Stage 1 of the methodology disaggregates the power sectors by fuel into power sectors by load-type (i.e. base and peak load). The base-peak split in this stage could be improved or, given data, these could split into distinct technologies (e.g. steam, combined-cycle, combustion turbine). The latter case would give much better estimates of cost structures as well as allow for more detailed modeling.
2. The assumptions on the cost structure of transmission and distribution greatly influence the results for the generation technologies.
3. Additional coverage of levelized cost data would reduce deviation between the original data sets.
4. The levelized costs used in this version are for new generating capacity. In GTAP many countries have capital values much lower than those implied by the levelized costs and production data. This may be due to depreciated (old)

generating capacity in the country. Adjusting for this may bring estimates of levelized capital costs more in line with the GTAP data.

5. Coverage of production and input taxes for specific electricity technologies is currently limited.
6. As discussed in Section 4.6, the disaggregated data could be used as an additional data source for the GTAP ‘ely’ sector. This might help reduce the deviations between the bottom-up and top-down models.

By making the disaggregation method transparent and publicly-available, the intent is to continuously improve the method and foundational data via the social accounting and CGE research community.

Despite these limitations, the GTAP-Power database contains the most detailed representation of electricity technologies in any publicly-available global CGE database. Furthermore, the database is a freely available database extension with a subscription to the main GTAP database. As opposed to many other databases and models of this type, the disaggregation is fully documented and accompanied by an example version of the software used to produce the database. This allows and encourage continuous improvement on the assumptions, methodologies, and data in the research community.

An electricity-detailed CGE database is the foundation for researching electric power in the global economy. The following chapter describes a model which leverages this database to give a highly-detailed representation of the electricity sector in global economic modeling.

CHAPTER 5. A DETAILED REPRESENTATION OF THE ELECTRICITY SECTOR FOR ECONOMIC EQUILIBRIUM MODELS

This chapter formulates a detailed representation of the electricity sector conducive to partial and general equilibrium modeling that explicitly and endogenously captures capacity utilization, capacity expansion, and their interdependency. Independent and joint validations of these two interdependent mechanisms in a partial equilibrium setting lend support to the predictive ability of the model. The validated electricity sector representation is also the most important feature of a general equilibrium version, termed the electricity-detailed general equilibrium (EDGE) model; although this particular extension is not discussed in this dissertation.

Electricity generation, as opposed to capacity, is the relevant economic good in balancing supply and demand in the power sector. The magnitude and mix of generation from different technologies has important implications for long-run sustainability issues such as mitigating greenhouse gas emissions and moderating energy consumption. Changes in electricity generation result from two distinct economic mechanisms: i) construction of new capital, termed *capacity expansion*, or ii) increases or decreases in operation of existing capital, termed *capacity utilization*. Long-term returns on capital investment in electric power technologies drive expansion, while utilization is the substitution in response to prevailing economic conditions, especially fuel prices, which is also termed fuel-switching. The two mechanisms are interrelated in that capital rents partly depend on how much generation is produced per unit of capacity (i.e. capacity factor) and short-term factor utilization may be counterbalanced by long-term expansion. These joint mechanisms are often overlooked in long-term projections of the evolution of

electricity generation in both technologically-rich, partial equilibrium optimization (bottom-up) and globally consistent general equilibrium (top-down) models.

Bottom-up models are technologically-rich and use exogenous projections of capital, fuel, operating, maintenance, and other costs to drive changes in total capacity of electricity generating technologies. Capacity factors for each technology also tend to be treated exogenously implying that existing capacity is unable to adjust to prevailing economic conditions, and new capacity must operate at a preordained level. One major criticism is that price projections are exogenous despite the possibility that both demands for fuel and their corresponding prices are partly determined by changes in the electricity sector. Endogenously determined capacity factors suggest the need for endogenously determined prices for the technologies.

Top-down, namely CGE, models have endogenously determined prices, but typically characterize the substitution of technologies in generation (GWh) as a single mechanism, which treats factor utilization and capacity expansions implicitly. This ignores a key distinction in how different economic and policy shocks impact the electric power sector. For instance, fuel price shocks (e.g. decline in gas prices as a result of the shale boom) can be adjusted for in the short-term with existing infrastructure, prior to construction of new capital and decommissioning of old. On the other hand, subsidies to capital (e.g. US solar and wind investment tax credits) impact only new capital investments.

In sum, bottom-up models capture technology-level, extensive and sector-level, intensive margins given sector-level, extensive margins, but often fail to capture complete market price responses; although several models include supply responses for input prices. Top-down models focus on the sector-level in terms of generation, but blur the lines between intensive and extensive margins at the technology-level (See Table 5.1).

The question addressed here is how to represent capacity utilization and expansion that endogenously respond to economic conditions in a manner consistent with existing partial and general equilibrium modeling frameworks. This chapter

Table 5.1.: Intensive and extensive margins for electric power at technology and sector-level

	Intensive margin	Extensive margin
Technology-level	Capacity factor utilization	Capacity expansion
Sector-level	Technological substitution	Total generation expansion

formulates a partial equilibrium model that explicitly and endogenously captures these interdependent mechanisms. A corollary contribution is a novel implementation of a variant of the constant elasticity of substitution function which ensures aggregate output quantity (i.e. total GWh) is equal to the sum of input quantities (i.e. individual GWh for each technology) for the electric power sector (van der Mensbrugge and Peters, 2015). Model estimates for annual capacity utilization, given exogenous capacity growth are validated against observations between 2002 and 2012. Next, a more complete validation exercise is undertaken combining both capacity utilization and endogenous expansion from 2007 to 2018. These validation exercises, often overlooked in large-scale modeling, demonstrate the predictive power of the electricity sector representation in a equilibrium modeling with capacity utilization, expansion, and their interdependency.

Identical to the sectors in GTAP-Power, the electricity generating technologies explored in this work are: nuclear, coal, gas, oil, hydro, wind, solar, and other (which is comprised of biofuels, geothermal, and other less-prevalent technologies). Gas and oil power is separated into base and peak load because combined-cycle plants are competitive in providing base load power, and combustion turbine plants are competitive in providing peak load power. Similarly, hydroelectric plants are split in to base and peak load because they can potentially be used for either or both types of load provision. Base and peak load technologies are designated by a “BL” and “P” suffix, respectively.

Instead of including the full general equilibrium linkages, the following sections describe a partial equilibrium model of the US electricity sector, which uses the GTAP-Power data with equations that are conducive to general equilibrium modeling. The complete economy-wide, general equilibrium supply schedule is replaced by upward supply curves for coal, oil, gas, and O&M, and the demand schedule is replaced by a downward sloping demand curve (see Figure 5.1). In the model prices are shifted to replicate observations and run simulations, but supply and demand are still price responsive and can capture rebound effects. Also, household income is exogenous (i.e. not impacted by electricity prices and tax), and there is no international trade. These simplifications allow for a controlled analysis and validation of the US electricity sector model – the relevant sector in this dissertation. The analysis lends credibility to the US electricity sector results as part of a complete general equilibrium analysis.







GTAP CGE model		Electricity-detailed computational equilibrium model														
	PROD_COMM 1	TnD	Nuclear	Coal	GasBL	OilBL	HydroBL	Wind	Other	GasP	OilP	HydroP	Solar	PROD_COMM n	HOU	GOV
coa		↑														
oil		↑														
p_c																
gas		↑														
gdt																
Ely																
O&M 1																
		↑														
O&M n																
Capital		↑														
PTAX																

Figure 5.1.: The white part of the matrix represents the partial equilibrium model described in Chapter 5 with supply curves for coal, gas, oil, O&M, and capital. The light gray areas signify the full GTAP CGE linkages which have greater sectoral detail as well as supply and demand schedules across all sectors, households, and government rather than simple supply curves. Integrating the electricity-detail in GTAP requires introducing the full sectoral detail for inputs as well as the demand for electricity for each user (i.e. firms, household, and government)

5.1 Conceptual Framework

The development of the model focuses on the following accounting relationship:

$$q_t^g = c_t \cdot \alpha \cdot q_t^c \quad (5.1)$$

where q_t^g is the quantity of GWh generated by technology t , c_t is the capacity factor (i.e. ratio of actual production to capacity) for technology t , q_t^c is the quantity of capacity (in GW), and α is the number of hours (one calendar year $\cong 8,760$ hours).

Equation 5.1 can be log-linearized (in percentage change terms) as follows:

$$\widehat{q}_t^g = \widehat{c}_t + \widehat{q}_t^c \quad (5.2)$$

where \widehat{q}_t^g is the percentage change (henceforth denoted by the lowercase letter and hat) in generation, \widehat{c}_t is the percentage change in capacity factor, and \widehat{q}_t^c is the percentage change in total capacity. Therefore, the quantity of generation from each technology can be increased (decreased) by increasing (decreasing) the capacity factor and/or the total capacity. The representation explicitly represents capacity utilization (\widehat{c}_t) and expansion (\widehat{q}_t^c) in the generation of electricity.

The associated equations, which describe the economic equilibrium model, are also written in percentage change form following the ORANI tradition (Dixon, 1982), and the non-linear model is solved as such using the GEMPACK software (Harrison et al., 2014). Level variables are not accented (as in equation 5.1), percentage change variables are written in lowercase with hat accents (e.g. \widehat{q}_t^g as in equation 5.2), and aggregate variables are written in uppercase (e.g. level and percentage change in *total* generation would be Q^g and \widehat{Q}^g , respectively).

5.2 Electricity Production

The production of electricity from existing capacity follows the standard GTAP production structure. Intermediate inputs (e.g. fuels) and an aggregate value-added,

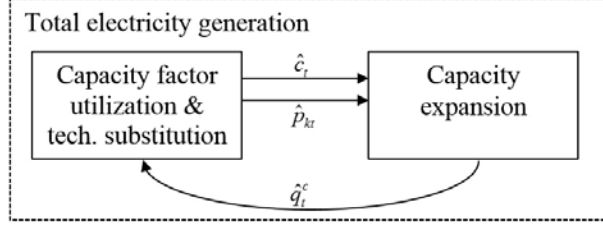


Figure 5.2.: Total electricity generation is determined by the dual capacity utilization and capacity expansion mechanisms and their interdependency. The two mechanisms are linked by the capacity factor for existing and new capacity (\hat{c}_t), returns to capacity ($\hat{p}_{k,t}$), and the net change in capacity (\hat{q}_t^c).

comprising capital and operating and maintenance (O&M), are Leontief inputs to production. Capital and O&M can substitute in the value-added nest, specified by a cost-minimizing constant elasticity of substitution (CES) parameter, σ_t^{va} . This behavior is described by the following equations:

$$\hat{q}_{it} = -\hat{a}_{it} + \hat{c}_t \quad \forall i \in INT \quad (5.3)$$

$$\hat{q}_{it} = -\hat{a}_{it} + \hat{c}_t - \sigma_t^{va} \cdot [\hat{p}_{it} - \hat{a}_{it} - \sum_{i \in VA} s_{it}^{va} \cdot (\hat{p}_{it} - \hat{a}_{it})] \quad \forall i \in VA \quad (5.4)$$

where \hat{q}_{it} is the percentage change in demand for input i by technology t , \hat{p}_{it} is the percentage change in price of input i in technology t , and \hat{a}_{it} is the percentage change in technological efficiency of input i in technology t . The variable s_{it}^{va} is the value share of input i in total value-added for technology t . The sets INT and VA contain the intermediate and value-added inputs, respectively.

5.3 Capacity Factor Utilization

The annual capacity factor is the ratio of annually generated electricity to the available capacity in a year. Since demand fluctuates on a daily and seasonal basis with dynamic marginal values of electricity, the annual capacity factor is an aggregate measure of the technology's annual supply response. That is, some technologies might

have a large capacity factor for a brief period during peak hours, but the annual capacity factor might be low since it is not a competitive technology during normal hours (e.g. oil power). Base load technologies tend to have higher annual capacity factors than technologies that only operate during peak loads.

The possibilities for capacity utilization adjustment is based on two determinants: i) flexibility of the technology itself and ii) substitution between flexible technologies. Flexibility is the ability of a technology to increase and/or decrease operations of the plant in response to the prevailing economic conditions, characterized by technical specifications of technology. Although flexible technologies can quickly increase or decrease operations, they are *not* perfectly substitutable due to factors such as space, time, and contract lead time (Hirth et al., 2014), so their substitutability is estimated and validated against historical observations.

Flexibility and substitutability of these technologies are discussed conceptually and integrated into the empirical model in their respective sections, Section 5.3.1 and Section 5.3.2. This work focuses on annual factor utilization, but both the conceptual and empirical components can be applied to smaller time intervals without loss in generality, provided appropriate data exist.

5.3.1 Capacity Factor Flexibility

The capacity factor can adjust to prevailing economic conditions only if the existing capacity of a technology is flexible in meeting demand. The capacity factor can be derived by dividing generation by the capacity-hours reported for a year.

Figure 5.3 shows annual capacity factors have changed over time in response to overall demand and prevailing economic conditions.

A flexible technology has the ability to adjust generation levels from built capacity - shown by a slope different than zero (i.e. Coal, Oil, GasBL, and GasP). GasP shows only slight flexibility because it comprises a large percentage of peak power. Therefore, small changes in capacity factor are large changes in generation; oil and

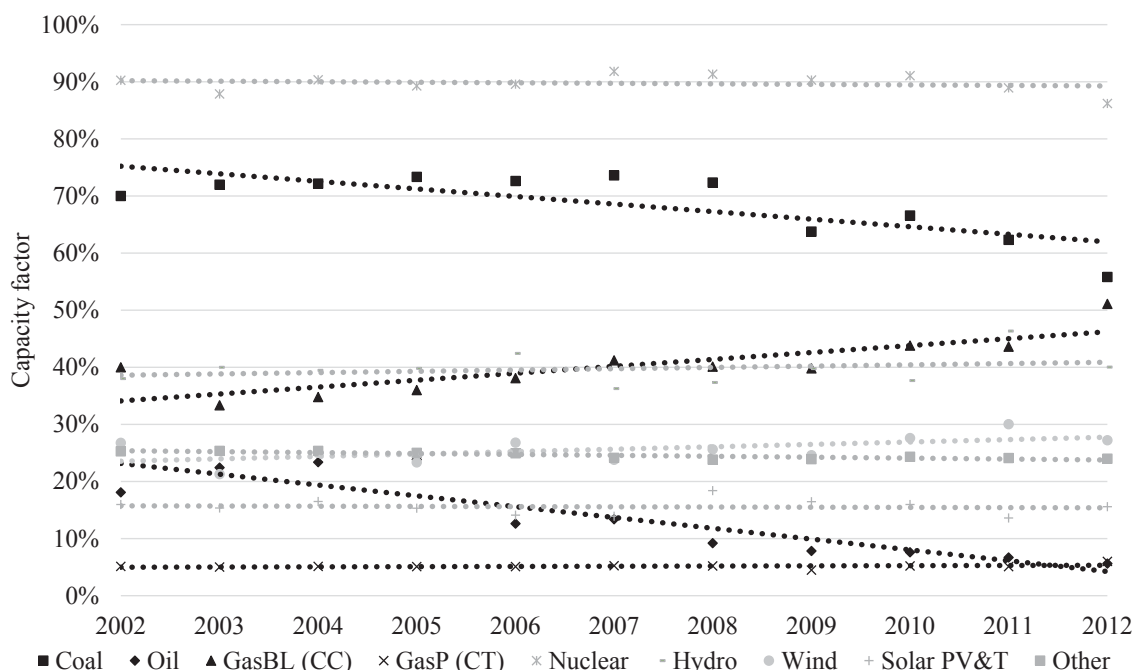


Figure 5.3.: Annual capacity factors from 2002–2012 EIA (2015b). The slope of the trend lines indicate the technology’s flexibility in the face of changing economic conditions over time. Black lines represent the flexible technologies, while the gray trend lines represent the inflexible technologies.

solar do not produce much in comparison. The inflexible technologies (i.e. Nuclear, Hydro, Wind, Solar, and Other) show little response to changing economic conditions. The only variability they exhibit is a result of normal annual operational fluctuations (e.g. plant shutdowns, maintenance); annual rainfall in the case of hydro power; wind in the case of wind power; and sunlight in the case of solar power. These types of non-economic fluctuations are not captured in this model.

The capacity factor can be represented by supply curves (for each technology) with capacity factor on the x-axis and price of technology-specific capital on the y-axis (Figure 5.4). As the capacity factor expands, it becomes increasingly costly to produce each unit of electricity. The increase in the cost of electricity is due to expensive short-term capital improvements. It is infinitely expensive to operate at full capacity (i.e. capacity factor of 100%) due to the non-economic requirements for plant maintenance. Flexible technologies (i.e. coal, gas, and oil power) can

adjust production without much additional capital expenditure (elastic supply) by increasing O&M inputs. Inflexible technologies cannot adjust production except with exceedingly large capital expenditure (inelastic supply).

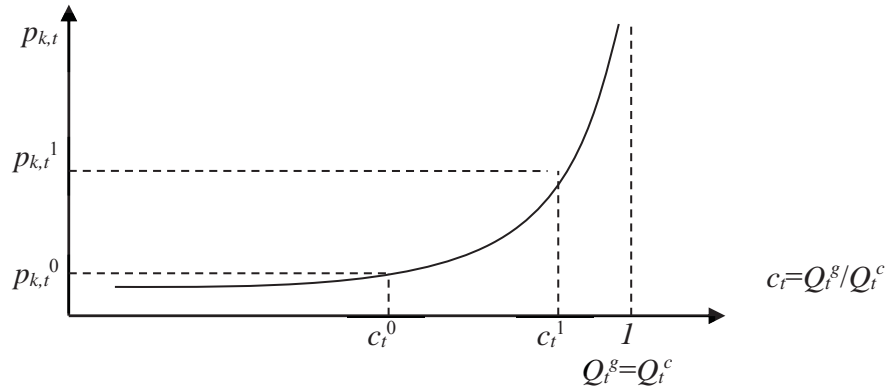


Figure 5.4.: Generic supply curve for a flexible electricity generating technology. An inflexible technology would be a vertical line intercepting the x-axis at the capacity factor value.

Supply for inputs to electricity production are represented in the model by simple elasticities of supply. Intermediate inputs and O&M are mobile inputs, but capital is technology-specific. These equations can be written as:

$$\hat{q}_i = \mu_i^s \cdot \hat{p}_i \quad \forall i \in MOBL \quad (5.5)$$

$$\hat{q}_i = \hat{q}_{it} \quad \text{and} \quad \hat{p}_i = \hat{p}_{it} \quad \forall i \in MOBL \quad (5.6)$$

$$\hat{q}_{it} = \mu_{it}^s \cdot \hat{p}_{it} + \hat{q}_t^c \quad \forall i \in NMOB \quad (5.7)$$

where μ_i^s and μ_{it}^s are the supply elasticities for mobile (*MOBL*) and non-mobile, technology-specific (*NMOB*) inputs, respectively. The set *MOBL* includes intermediate inputs and O&M while set *NMOB* is only for capital, which is fixed in the short-term (i.e. $\mu_{it}^s = 0$). Chetty et al. (2011) finds a labor supply elasticity of 0.30, which is used for O&M. Brown (1998) finds long-run supply elasticities of coal, oil, and gas to be 1.86, 0.51, and 0.76, respectively. These estimates are pre-shale oil

and gas boom, and we might expect the supply of oil and gas to be more elastic due to the ability to drill shale wells much more quickly than conventional, capital-intensive wells. The sensitivity of the model results to these parameters can be tested using systematic sensitivity analysis.

Flexible technologies can adjust capacity utilization by substituting O&M for capital improvements. For example, plant operators can adjust labor hours (e.g. paying overtime) and the frequency of normally scheduled maintenance. The parameter σ_t^{va} (in equation 5.4) is calibrated such that the returns to capital roughly match the change in capacity factor. That is, the returns reflect the change in production (i.e. generation) per unit of capital (i.e. capacity). Figure 5.5 shows the capacity factor supply curves for flexible technologies which are drawn out by shifting demand for the technology. A higher CES parameter between value-added corresponds with greater flexibility.

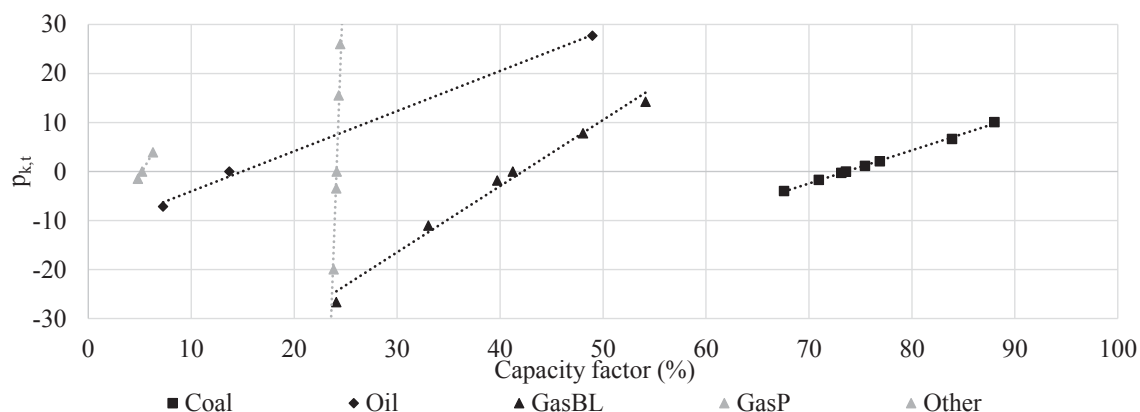


Figure 5.5.: Capacity factor supply curves for flexible electricity generating technologies. Data points represent results from shocks to other substitutable technologies to shift demand for the relevant technology. The selection of the shocks were designed to map out the response over a wide range of possible shocks.

The supply curves distinguish flexible and inflexible technologies. Inflexible technologies are excluded from Figure 5.5, but would be vertical lines at their current capacity factor. The parameters should ideally be estimated in conjunction with capacity expansion since they related to the long-run returns to capital. The

selection of parameters here (i.e. inelastic supply for inflexible technology and elastic supply for flexible) is sufficient for the capacity utilization to ensure that only flexible technologies respond to the prevailing economic conditions.

5.3.2 Substitution of Existing Capacity

Substitution of generation in total electricity production in both partial and general equilibrium modeling is often represented with CES parameters where generation from different technologies are directly substituted based on cost minimizing assumptions Paltsev et al. (2005); Wing (2006); Château et al. (2014). However, this representation treats the dual mechanisms of factor utilization and capacity expansion implicitly. Also, the standard CES production function does not preserve quantities. That is, due to the CES production specification, the derived generation from each individual technology may not sum to the aggregate generation of electricity. The CES assumes inputs are heterogeneous despite the fact that both are measured in GWh terms. The heterogeneous input assumption is well-suited for many problems, but can be problematic for interpreting the modeling results when corollary impacts flow from the quantity of electricity generation (e.g. emissions): how can they not sum? What is the correct total change in GWh? What is the impact?

Therefore, another contribution in this work, beyond the dual mechanism for changes in electricity generation, is a variant of the CES which preserves quantities (GWh) in the aggregate output and sum of disaggregate inputs, termed here as QCES. The variant has been implemented in the context of labor (Dixon and Rimmer, 2003) and land (Giesecke et al., 2013) markets and requires a slightly different conceptual justification from the standard CES. This section discusses the conceptual justification and implements the QCES specification for capacity factor adjustment in the electricity sector.

5.3.2.1 Standard CES Substitution

CES is the most common specification for the supply of electricity in both partial and general equilibrium models (Paltsev et al., 2005; Wing, 2006; Château et al., 2014). Electricity generating technologies substitute with one another in some form of nested CES where the nesting attempts to control for some of the operational considerations of the electricity sector (e.g. load type). Substitution is based on cost minimization with a CES production constraint. The formulation for a single nest is as follows:

$$\min_{q_t^g} C = \sum_t p_t^g \cdot q_t^g \quad (5.8)$$

subject to:

$$Q^g = \left[\sum_t (\theta_t \cdot q_t^g)^\rho \right]^{\frac{1}{\rho}} \quad (5.9)$$

$$\sigma = \frac{1}{\rho - 1} \quad (5.10)$$

where C is the cost function characterized by the sum of the cost per unit production from each technology, p_t^g , multiplied by the total production of that technology, q_t^g . Total production of electricity, Q^g , is characterized by CES production where θ_t are the share parameters, and ρ is the CES exponent, which is easily transformed to the familiar CES parameter, σ , describing the constant elasticity of substitution amongst inputs.

Implemented in log-linearized form, derived demands for each technology are given by the following equation (Dixon, 1982; McDougall, 1992):

$$\widehat{q}_t^g = \widehat{Q}^g - \sigma \cdot \left(\widehat{p}_t^g - \sum_t s_t^v \cdot \widehat{p}_t \right) \quad (5.11)$$

where s_t^v is the value share of technology t in the total production nest. Again, the hat accent refers to the percentage change in the variables. The standard CES specification is attractive for both partial and general equilibrium analysis for several reasons. First of all, while the assumed CES production function is largely

arbitrary (outside of the fact other common production specifications, such as Leontief and Cobb-Douglas, are special cases of the more general CES specification), the cost minimization objective seems appropriate for electricity aggregators choosing which technologies to employ. Also, only a single parameter, σ , is required to characterize substitution (s_t^v is given by the data). This point is important for calibrating large-scale models which already require a vast number of equations and corresponding parameters (i.e. supply, demand, and substitution parameters).

However, the standard CES function does not preserve additivity of the inputs in the output. This point is harmless in the case of transforming fundamentally different goods into a new good (e.g. capital and labor) or in the case of different qualities across inputs where the units of the aggregate and the inputs are identical (e.g. quality-adjusted labor (Bowles, 1970)). However, this is not the case for electricity in which units of production for each technology and the units of the aggregate are identical and of the same quality.¹

This motivates the implementation of the QCES described by Dixon and Rimmer (2002, 2003, 2006) below; however, QCES first requires a slightly modified theory regarding the structure of the electricity sector than the standard CES (van der Mensbrugghe and Peters, 2015).

5.3.2.2 Quantity-Preserving CES (QCES) Substitution

Supply must equal demand instantaneously in an electricity network. As a result, the values of electricity produced with different generation technologies (from the system operator perspective) are not stationary over any period of time due to: i) the nature of demand, which can fluctuate by the minute, hour, day, and season and ii) the operational constraints of technologies that may prevent flexibility in responding to the fluctuating demand. While it may be intuitive to think of cost minimization as

¹Although values of electricity produced from different technologies may be heterogeneous in supply due to time, space, and contract lead-time (Hirth et al., 2014), the qualities of the produced electricity from each are identical.

in the standard CES, the average cost of producing electricity is an incomplete idea in that it does not reflect the complete cost to the system operator because of costs incurred by operational constraints, which are difficult to identify and incorporate into the market prices. For instance, the Pennsylvania-New Jersey-Maryland (PJM) market in the United States refers to these costs as: day-ahead and balancing operating reserves, reactive services, synchronous condensing, and black start services. The purpose of identifying and categorizing these costs is “to reflect the impact of physical constraints in market prices to the maximum extent possible (Monitoring Analytics, 2015, p.g. 142)” mainly for the purposes of reliability of instantaneously adjusting supply (and having reserves) to meet the unpredictable demand (Monitoring Analytics, 2015).

Recognizing that average costs of generation are incomplete, there is an unobserved *utility of supply* that reflects of the usefulness of the supply from different technologies in meeting the specific nature of the demand which balances of average cost of generation, reliability costs, and costs incurred from operational constraints.

Therefore, the problem of the system operator is two-fold. First, the total production from each technology, q_t^g , must meet the total electricity demand observed over some period of time, Q^g . Second, the operator maximizes the utility of the supply where the utility is defined as a CES function with revenue obtained from each generation technology. That is, the system operator gets some positive contribution to overall utility from revenue obtained from certain technologies. Here, revenue, rather than the technologies, is substitutable.

The QCES specification maximizes the CES utility subject to the sum of electricity produced from each technology equaling the total electricity demand, as shown below:

$$\max_{q_t^g} U = \left[\sum_t (p_t^g \cdot q_t^g)^\rho \right]^{\frac{1}{\rho}} \quad (5.12)$$

subject to:

$$Q^g = \sum_t q_t^g \quad (5.13)$$

$$\sigma = \frac{1}{1 - \rho} \quad (5.14)$$

The observed mix of revenue from various electricity generation technologies is then the mix which optimally satisfies the complex nature of demand given by the unobserved utility. Note also that the QCES parameter is related to the CES exponent in a slightly different manner than in the standard CES case (Equation 5.10).

The log-linearized description is nearly identical to the standard CES specification (equation 5.11) except the price index is based on the quantity shares, s_t^q , rather than the value shares s_t^v (van der Mensbrugghe and Peters, 2015).

$$\hat{q}_t^g = \hat{Q}^g - \sigma \cdot \left(\hat{p}_t^g - \sum_t s_t^q \cdot \hat{p}_t \right) \quad (5.15)$$

Because s_t^q comes from observable data, implementation of the QCES into partial and general equilibrium models is straightforward. Perfect substitution, $\sigma = \infty$, would imply that a small increase in revenue for a unit of utility (e.g. price or quantity per “util” decreases) in a technology would result in all electricity being produced from that technology, just as we would observe in the traditional CES with production as inputs. Leontief, $\sigma = 0$, implies that the same proportion of revenue is required for a change in utility or change in total electricity since utility is homothetic in quantities.

A nested CES-type structure is commonly used in CGE analysis of electricity production. The motivation of this work to incorporate economy-wide linkages, so a commensurate representation is used here. The important consideration in creating a reasonable the production structure is capturing imperfect substitution between technologies, especially regarding the dispatchability and the load type (base versus peak). Château et al. (2014) does not include electricity technologies. Instead, capital and fuels are imperfectly substitutable, coal and liquids (i.e. crude oil, refined oil, and gas) are imperfectly substitutable, and liquids are imperfectly substitutable. McDougall and Golub (2008) characterizes a similar nested substitution between

fuels and between fuels and capital. However, both of these ignore non-fuel-based technologies in electricity production. Pant (2007) uses a technology bundle approach with specific generating technologies, but does not consider any dispatchability and base versus peak load characteristics in the substitution between the technologies. Paltsev et al. (2005) uses a nesting between non-dispatchable and dispatchable (i.e. fossil fuels) and then substitution between the dispatchable fossil fuels. Variable technologies (i.e. solar and wind) are accompanied by a peak-load dispatchable technology (e.g. gas or oil power). The parameters in the nesting structure are neither estimated nor calibrated.

The capacity utilization model here uses a nested QCES production structure to represent substitution between technologies. Capacity factor utilization is represented as a nested derived demand system for the technologies represented in Figure 5.6. Base and peak load technologies are distinguished from one another to reflect the additional operational aspect of the electricity sector. QCES parameters are calibrated across several time periods for base and peak load (i.e. σ^{bl} and σ^{pl} , respectively) separately. The substitution between base and peak load (i.e. σ^g) and between transmission and distribution (T&D) and total generation are assumed equal to zero (i.e. Leontief).

5.3.2.3 Calibrating the Quantity-Preserving CES Parameters

The QCES parameters for the United States are calibrated across several time periods (i.e. 2002–2012) for the base and peak load nests in Figure 5.6 from equations 5.16 and 5.17, respectively, where the percentage change variables are for annual time steps. The error terms, ϵ_{bl} and ϵ_{pl} , are minimized independently using an ordinary least squares estimator.

$$\hat{q}_t^g = \hat{Q}^g - \sigma^{bl} \cdot \left(\hat{p}_t^g - \sum_t s_t^q \cdot \hat{p}_t \right) + \epsilon_{bl} \text{ for } t \in EBL \quad (5.16)$$

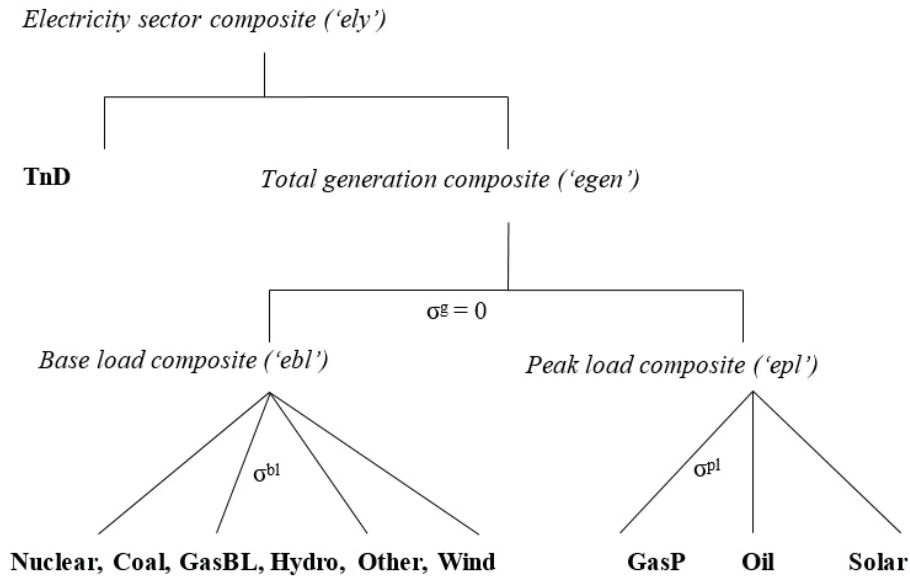


Figure 5.6.: Production structure for capacity utilization. Composite sectors are in *italics*.

$$\hat{q}_t^g = \hat{Q}^g - \sigma^{pl} \cdot \left(\hat{p}_t^g - \sum_t s_t^q \cdot \hat{p}_t \right) + \epsilon_{pl} \text{ for } t \in EPL \quad (5.17)$$

where *EBL* and *EPL* are the sets of base and peak load technologies, respectively. Therefore, the relevant data are annual generation for each technology, q_t^g , and annual costs of generation by technology, p_t^g . The variable Q^g can be constructed by summing q_t^g across all technologies, and s_t^q is constructed by share-weighting q_t^g over the *EBL* and *EPL* sets. EIA (2015b) and other EIA databases have annual data on q_t^g from 2002 to 2012. Moving from annual level variables to annual percentage change variables is straightforward.

Changes in annual costs of generation by technology are constructed from the following equation:

$$\hat{p}_t^g = \sum_i c_{it} \cdot \hat{p}_{it} \quad (5.18)$$

where c_{it} is the share of the cost of input i (i.e. fuel, O&M, and capital) in producing using technology t (i.e. cost structure), and \hat{p}_{it} is the percentage change in price of input i in technology t . Cost structures for each technology are taken from the

GTAP-Power database (Chapter 3). Annual values for p_{it} are required for each technology to determine p_t^g ; alternatively, annual *changes* in price p_{it} from the base levelized cost estimates in Chapter 2 are required to determine p_t^g . BLS (2015) has data on labor costs, $p_{O\&M,t}$, in the total electric power sector, and uniform changes, $\Delta p_{O\&M,t}$, are applied across all technologies. The impact on total price will depend on the share of O&M in the cost structure of the particular technology. Fuel prices for coal, oil, and gas (measured in real price per MMBTu observed in the electricity sector) are also available from EIA (2015b) for the relevant time period. Capital costs are assumed to be constant since we are investigating short-term changes in generation from factor utilization only ($\hat{p}_{k,t} = 0$).

The multi-period calibration procedure estimates annual QCES parameters for base and peak load technology substitution to be 0.462 and 0.472, respectively. More observations would be needed to econometrically estimate these parameters (here $n = 10$). It is worth noting that general equilibrium modeling takes the perspective that prices and quantities are endogenously determined, as opposed to exogenous prices in this estimation, and the greater concern are the feedbacks in the structure of the economy. It may be worth varying the parameter in the CGE model using systematic sensitivity analysis to represent the uncertainty (Arndt, 1996). For example, an increase in substitution would make utilization more elastic. More observations would be required to be certain of the “true” parameter value (if a “true” parameter value even exists).

Of greater interest than the significance of the parameter values is how well these parameter values characterize factor utilization in an economic model. While the lack of econometrically supported parameters in these models is one of the major criticisms of DeCanio (2003), the detailed representation of electricity and the validation exercises addresses two other major criticisms. The following section describes the capacity utilization validation.

5.3.3 Capacity Factor Utilization Validation

The parameters mapping the flexibility of supply from existing capacity, σ_t^{va} , and parameters describing substitution between existing capacity, σ^{bl} and σ^{pl} , are now specified in the capacity utilization version of the model with exogenously specified capacity expansion (i.e. expansion is controlled in these simulations). The model is subjected to shifts to the main determinants of factor utilization (i.e. fuel prices) and some important longer-term determinants of demand and supply (i.e. per capita income, population, and technical efficiency of the electricity sector). The following total electricity demand equation and zero-profit condition, combined with derived demands and supply response for inputs (Equations 5.3–5.7) and derived demands for electricity generating technologies (Equations 5.16–5.18), complete the equilibrium model for capacity utilization:

$$\widehat{Q}^g = -\widehat{a}^g + \widehat{y} + \widehat{pop} + \sum_u s_u \cdot \mu_u \cdot \widehat{P}^g \quad (5.19)$$

$$\widehat{P}^g = \sum_t s_t^g \cdot \widehat{p}_t^g \quad (5.20)$$

where \widehat{a}^g is the percentage change in technical efficiency of electricity use, \widehat{y} is the percentage change in gross national income per capita, \widehat{pop} is the percentage change in population, s_u is the end-use share of total electricity demanded by user category u (comprised of industrial, commercial, and residential), μ_u is the demand elasticity for electricity by user, and \widehat{P}^g is the aggregate price of demanded electricity.

5.3.3.1 Baseline Factor Utilization Validation

The model validation uses a base year of 2007 from which the power sector is shifted retrospectively to 2002 and prospectively to 2012.² The results for the change in capacity factor for each technology are then compared to the observed capacity

²The baseline validation was initially performed with the GTAPv8 database with a base year of 2007. Using a base year of 2007 is advantageous because it allows for validation before and after structural change in fuel prices brought on by the shale oil and gas boom in the US.

factors over that same time period. The deviations from the observed values are expected to become more pronounced as the deviation from 2007 increases because of the compounding influence of other uncontrolled longer-term determinants as well as the interaction between existing and new capacity (which is introduced in the full model later). Still, the capacity utilization validation lends confidence in the model's ability to capture the response of existing electricity generation technologies in the aggregate region to economic and policy shocks.

Table 5.2.: Shocks to key drivers of capacity factor from 2002–2012. 2007 is the reference year for GTAPv8. In percentage change from reference year. All are exogenous shifts except \hat{a}^g , which is an output from the model.

	Exogenous shifts (percentage deviation from 2007)								Model Output
Year	\hat{y}	\hat{pop}	\hat{p}_{coal}	\hat{p}_{gas}	\hat{p}_{oil}	$\hat{p}_{O\&M}$	\hat{Q}^g	\hat{q}_t^c	\hat{a}^g
2002	-9.844	-4.517	-18.622	-42.304	-46.322	7.177	-7.176		2.290
2003	-7.612	-3.693	-18.526	-14.592	-31.962	4.250	-6.581		3.550
2004	-4.863	-2.797	-15.680	-8.009	-34.339	5.870	-4.479		3.230
2005	-2.602	-1.897	-7.648	22.566	-4.662	-2.729	-2.438		1.310
2006	-1.181	-0.947	-1.820	0.369	-10.653	0.366	-2.214		1.930
2007	-	-	-	-	-	-	-		-
2008	-0.947	0.950	12.602	22.148	45.969	2.442	-0.899	EIA (2015)	-0.756
2009	-4.639	1.839	20.647	-35.582	-5.395	6.760	-4.966		4.900
2010	-2.728	2.695	21.922	-31.942	26.491	10.082	-0.762		1.310
2011	-1.723	3.483	24.440	-38.820	60.409	13.766	-1.362		2.750
2012	-0.384	4.276	21.407	-56.569	57.157	14.019	-2.622		7.060

Table 5.2 shows the exogenous shifts imposed in the model. Population consistently increases from 2002 to 2012. Coal prices increase from 2003 to 2008 and are relatively flat otherwise. Gas prices increase from 2002 to 2009 then fall sharply as the US shale oil and gas boom, which began in 2007, takes shape. Oil prices generally increase over the time period but with large fluctuations. Fuel prices per MWh of electricity generated are shown in Figure 5.7.

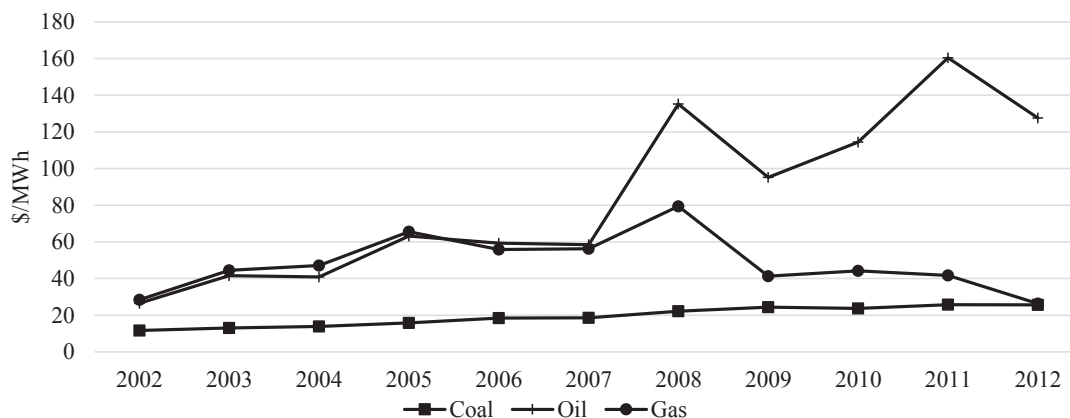


Figure 5.7.: Fuel prices per MWh of electricity produced (nominal dollars)

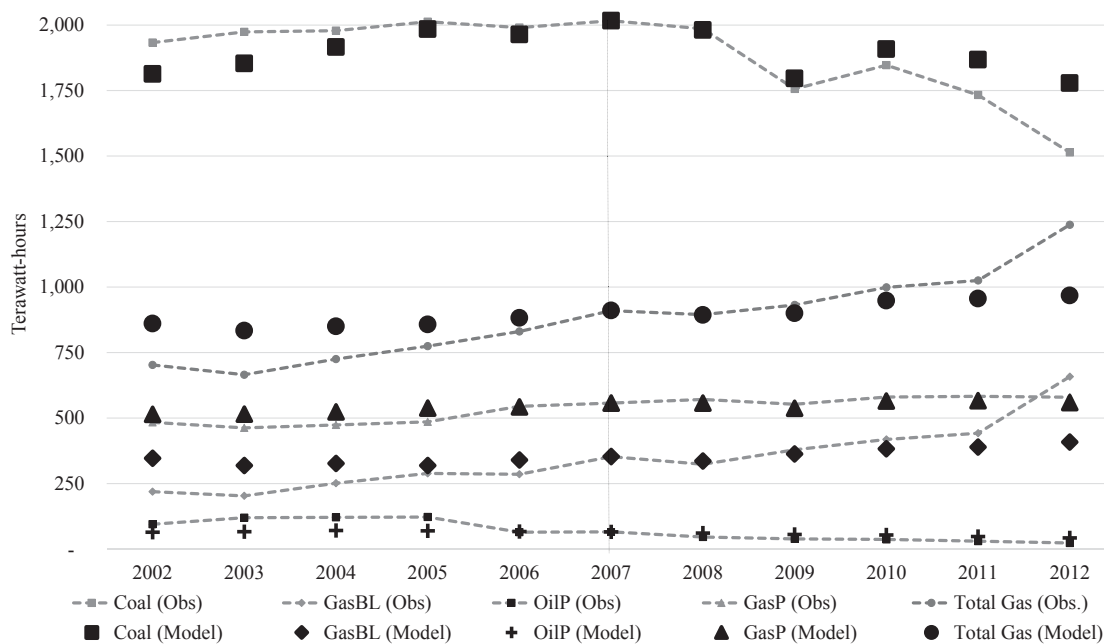


Figure 5.8.: Validation of capacity factor portion of model. Model results are represented by large markers and are not connected because they are each shifted separately from the 2007 base year. Observed values are gray, dashed lines.

The model results correlate well with the observed values (see Figure 5.8 and Table 5.4)³; however, there are two important time periods with larger deviations

³Model results are represented by large markers and are not connected because they are each shifted separately from the 2007 base year. That is, 2007 to 2008, 2007 to 2009, etc. This representation is used in the charts throughout.

between the observed generation and the model results. First, between 2002 and 2005 the model projects too little coal generation and too much gas generation in base load power, and the model projects too little oil generation in peak load power. Second, between 2011 and 2012 the model projects too much coal generation and too little gas generation in base load. This may be partly due to attempts by the federal government to regulate mercury as well as structural change in coal contract duration (leading to increasing substitutability from coal with time).

In March 2005, the Clean Air Mercury Rule (CAMR) issued by the US Environmental Protection Agency (EPA) sought to permanently cap and reduce mercury emissions from power plants by 70%. This was later vacated in 2008, but it signaled that the federal government was targeting mercury emissions by coal and oil-fired generation. In the context of the model, prior to 2005 the actual cost (excluding the cost of prospective regulation) of coal and oil power may actually be lower than the shifts imposed on the 2007 economy.⁴ This would explain some of the deviation between the observed and model results between 2002 and 2005.

Subsequently in March 2011, the EPA announced the Mercury and Air Toxics Standards (MATS) which aimed to reduce mercury emissions from coal and oil emissions more aggressively by up to 91%. This standard would contribute to a higher actual cost (including prospective regulation) of coal and oil power than in the model shifts after 2011.⁵ This might explain part of the deviation in years 2011 and 2012.

While neither the CAMR nor MATS is law at the time of writing this dissertation, several utilities have already installed equipment to limit mercury emissions in anticipation of such legislation (EIA, 2014). The regulations also may have signaled to the electric power industry that they should have higher future cost expectations for coal-fired generation; this permeates the validation.

⁴The projected annualized compliance cost of CAMR and CAIR (a related measure for sulfur oxide and nitrous oxide) is \$2.5 billion or about 1.5 \$/MWh (USD 2007) (EPA, 2005)

⁵The projected annualized compliance cost of MATS is \$9 billion for coal-fired generation or about 4.7 \$/MWh (USD 2007) (EPA, 2011)

In addition, the elasticities of substitution used in the validation were calibrated on an annual basis to control for interactions with capacity expansion. However, the long-term elasticity of substitution in base load power might be greater because of the decreasing term of coal contracts. In the US electricity sector 93% of coal in 2011 was purchased on contracts longer than one year with a median around three or four years which limits substitutability by contractually obligating coal generation. Only 44% of gas was purchased on long-term contracts in 2011 (Macmillan et al., 2013). The trajectory of observed and model estimates for gas and coal in base load power in Figure 5.8 seems to imply that, in fact, the two became more substitutable around 2010 which is about three years after the shale oil and gas boom, when natural gas prices dropped significantly compared to coal (Figure 5.7). The greater substitution observed after 2010 is partly a result of allowing long-term coal contracts to expire or, at least, allowing renegotiation due to a drop in natural gas prices. The elasticity of substitution will likely increase as longer-term coal contracts expire and are replaced by shorter-term contracts or even spot prices.

5.3.3.2 Policy-Adjusted Factor Utilization Validation

Top-down models are generally limited to price-based policies, and structural change is represented implicitly via elasticities (Hourcade et al., 2006). Both limitations make the aforementioned federal and industry policies a challenge to introduce in the model. Instead, monetary costs of meeting mercury regulations and a higher elasticity of substitution after the significant drop in gas prices in 2009 are used to adjust for the aforementioned federal and industry policies. These coarse adjustments should help to explain some of the gap between the baseline validation and the observed values.

Table 5.3 shows the new shifts which are used in addition to the shifts the validation described in Table 5.2 along with alternate elasticities of substitution for base load power to account for the reduction in coal contract duration. We assume

that the average coal contract duration is three years with expiration uniformly distributed over those three years.⁶ Although in reality contracts have different beginning and end points as well as duration, we assume that coal contracts will expire, shorten, or be significantly renegotiated beginning in 2008 as a result of the sharp decrease in gas price (i.e. one-third expire in 2008, one-third in 2009, and one-third in 2010). This implies that the estimated annual elasticity of substitution for base load power would double in 2009 and triple in 2010 (i.e. all the number of contracts now are expired or available for renegotiation) where it remains in the long-term.

Table 5.3.: Additional shifts to key drivers of capacity factor from 2002 - 2012. 2007 is the reference year for GTAPv8. In percent change from reference year. All are exogenous shifts except \hat{a}^g , which is an output from the model.

	Additional exogenous shifts			Model Output
Year	\hat{t}_{coalbl}^c	\hat{t}_{oilp}^c	σ^{bl}	\hat{a}^g
2002	-7.5	-7.5	0.462	2.570
2003	-7.5	-7.5	0.462	3.840
2004	-7.5	-7.5	0.462	3.520
2005	-7.5	-7.5	0.462	1.580
2006	-	-	0.462	1.930
2007	-	-	-	-
2008	-	-	0.462	-0.756
2009	-	-	0.924	4.350
2010	-	-	1.386	0.926
2011	-	-	1.386	2.270
2012	23.5	23.5	1.386	5.770

⁶An average of three years is consistent with volume-weighted state average contract lengths in Macmillan et al. (2013). Kozhevnikova and Lange (2009) report an average coal contract duration of 4.40 years between 1979 and 1999, but contract durations have been decreasing since that study period due to the fall in gas price (EIA, 2015c).

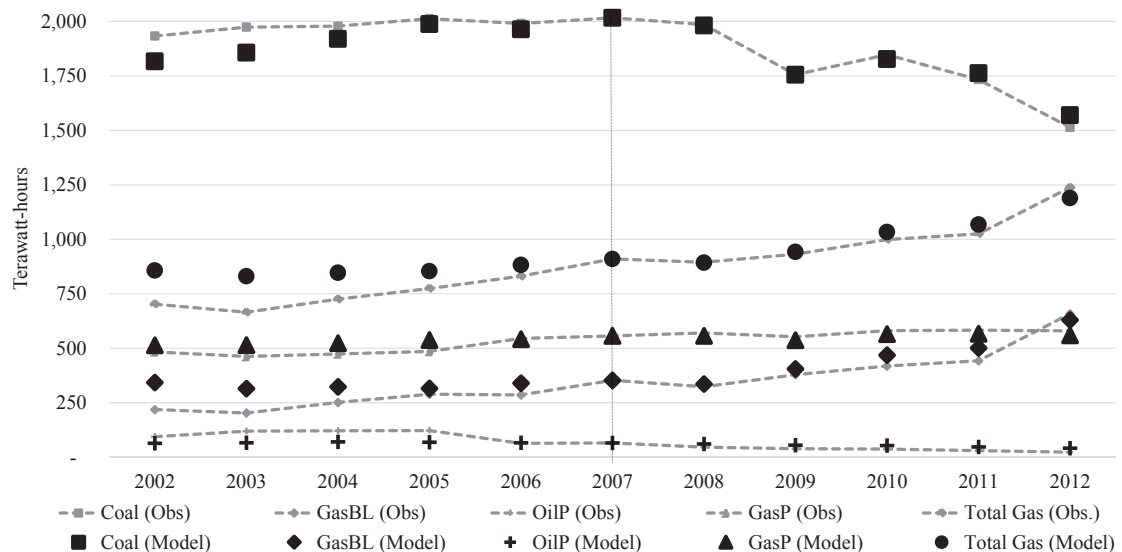


Figure 5.9.: Refinement of capacity factor validation using additional insights. Model results are represented by large markers and are not connected because they are each shifted separately from the 2007 base year. Observed values are gray, dashed lines

Figure 5.9 shows the results of the validation after the regulation and parameter adjustments. The numerical value of the refined shifts and parameters are crude estimates since the regulations were not passed, and there is insufficient data to estimate how the elasticity of substitution in the electricity sector may have changed in response to a decoupling of oil and gas prices and the subsequent fall in gas prices. However, these insights are shown to be an important aspect in modeling the evolution of electricity. Table 5.4 shows that these qualitative insights improve the correlation between the observed generation and the model predictions. Therefore, these retrospective insights are retained in the subsequent analysis.

There still remains a large difference in the predictions for coal and gas in base load and between oil and gas in peak load generation from 2002 to 2005. This may be indicative of the threat of future regulation of mercury in coal and oil power imposed a higher expected future cost of these generation types than that of just the expected upgrade costs to meet CAMR regulations. Coal has been a target for regulators for some time.

Table 5.4.: Comparison of correlation between observations and model predictions for the baseline and the policy-adjusted validation.

	Baseline	Policy-adjusted	Gain
CoalBL	0.743	0.947	+0.204
GasBL	0.899	0.950	+0.051
OilP	0.845	0.847	+0.002
GasP	0.923	0.924	+0.001
Total Gas	0.953	0.956	+0.003

5.3.4 Summary of Capacity Utilization

This section presented the capacity utilization portion of the electricity-detailed model and shows that it performs well with exogenous capacity. The substitution across generation types using existing capacity results in a change in the rates of returns to each technology. For instance, if the capacity factor of gas power increases (\hat{c}_t) more than other technologies, then gas power will have a higher return per unit of existing capacity ($\hat{p}_{k,t}$). This increases the value of the technology in the capacity expansion stage. Furthermore, a higher capacity factor in a flexible technology might crowd-out some of the returns from inflexible technologies. For instance, if gas power becomes inexpensive, nuclear power may see diminished returns on its capital. These rental rates partly determine the amount of investment and expansion in each type of capacity in the following section. Both the capacity factor and the rates of return from the utilization portion are passed on to the capacity expansion portion of the model, which is presented and analyzed in the following section.

5.4 Capacity Expansion

It is important to define the difference between nominal (or nameplate) capacity and effective generating capacity. Nominal capacity refers to the actual MW of capacity installed, while effective capacity refers to the capacity that can reasonably

be used to provide electricity generation (i.e. weighted by the capacity factor). Since generation is what balances supply and demand, effective capacity is important for decision-making in expansion. In the electricity-detailed model the percentage change in quantity of capital (i.e. \widehat{q}_t^c) directly maps to percentage changes in nominal capacity. Recall the motivating relationship in equation 5.2, the capacity factor is given as:

$$\widehat{c}_t = \widehat{q}_t^g - \widehat{q}_t^c \quad (5.21)$$

where \widehat{c}_t is the percentage change in capacity factor for technology t , \widehat{q}_t^g is the percentage change in generation by technology t , and \widehat{q}_t^c is the percentage change in nominal capacity for technology t . The relationship between nominal and effective capacity is given by:

$$q_t^e = q_t^c \cdot c_t \quad (5.22)$$

where q_t^e and q_t^c are the effective and nominal capacity, respectively. The total effective capacity changes can be decomposed into additional and retiring capacity. The rate of nominal retirements are a function of the changes in returns to capacity, \widehat{p}_t^c , and the annual rate of retirements defined by the inverse of the technical lifetime of existing plants. This captures the “economic lifetime” of the plant where a plant may extend its lifetime if the rate of return is higher or shorten if the returns become lower.⁷

$$\widehat{q}_t^{cr} = \frac{(100 - \widehat{p}_t^c)}{100} \cdot r_t \cdot t \quad (5.23)$$

$$\widehat{q}_t^{er} = \widehat{q}_t^{cr} \cdot cf_t \quad (5.24)$$

where \widehat{q}_t^{cr} and \widehat{q}_t^{er} are the percentage change in nominal and effective capacity for technology t due to retirements, respectively. The coefficient cf_t links the change in capacity utilization to the effective capacity ($cf_t \equiv \frac{\widehat{c}_t}{100} - 1$). This coefficient factor is

⁷Data prior to 2014 is given as net capacity changes rather than distinguishing between additions and retirements.

one of the three linkages between capacity utilization and capacity expansion shown in the original conceptual model in Figure 5.2.

Additional effective capacity is the sum of the net effective capacity and total retired effective capacity net of changes in total capacity utilization from all technologies.

$$\widehat{Q}^g - \widehat{C} = \widehat{Q}^{ea} - \sum_t s_t^q \cdot \widehat{q}_t^{er} \quad (5.25)$$

where \widehat{C} is the percentage change in total capacity factor and \widehat{Q}^{ea} is the required additional effective generation in the sector, and s_t^q is the share of electricity generated by technology t . This accounting condition is in generation (effective capacity) terms to ensure additional capacity meets the generation-based requirements.

We can determine the effective capacity additions for each technology according to the following equations:

$$s_t^q \cdot \widehat{q}_t^{ea} = s_t^a \cdot \widehat{Q}^{ea} \quad (5.26)$$

$$\widehat{q}_t^{ca} = cf_t \cdot \widehat{q}_t^{ea} \quad (5.27)$$

where \widehat{q}_t^{ea} and \widehat{q}_t^{ca} are the effective and nominal capacity additions, respectively. The coefficient s_t^a is the share of effective capacity additions allocated to each technology t .

These shares are derived using a multinomial logit (MNL) model where the utility U_t is solely a function of the change in rate of return on the capital in technology t , that is $U_t = \alpha \cdot P_t^c$ where the coefficient α , marginal impact on utility from rate of return, is assumed identical across generation types and can be calibrated to data.⁸

⁸This highly-constrained MNL model is validated later. Given additional data, an alternate MNL model could be estimated with econometrics

The variable P_t^c is the level of rate of return on new capacity and is linked to rental rates of existing capacity by the following equation.⁹

$$\widehat{p}_t^c = \widehat{p}_{k,t} + \widehat{a}_t - \widehat{t}_t^c \quad (5.28)$$

where \widehat{p}_t^c is the percentage change in rate of return of new capacity, $\widehat{p}_{k,t}$ is the percentage change in rental rate of existing capacity due to change in capacity factor which comes from the capacity utilization portion of the model, \widehat{a}_t is the percentage change of technological efficiency of new capacity (compared to existing capacity), and \widehat{t}_t^c are capital taxes/subsidies for new capacity.

This specification results in the following equation for s_t^a :

$$s_t^a = \frac{e^{\alpha \cdot P_t^c}}{\sum_t e^{\alpha \cdot P_t^c}} \quad (5.29)$$

Thus, the resulting long-term capital growth, capacity expansion, is:

$$\widehat{q}_t^c = \widehat{q}_t^{ca} - \widehat{q}_t^{cr} \quad (5.30)$$

This is the third linkage by which capacity expansion feeds back into capacity utilization for flexible technologies because net changes in capacity will impact the utilization rates of the sector.

5.5 Joint Capacity Utilization and Expansion Validation

The purpose of the joint capacity utilization and expansion validation is to test how the model performs in predicting both total capacity expansion as well as contributions from each generating technology given changes in capital rents from capacity utilization – as shown in the previous validation.

⁹Here it is assumed that the original rate of return is equal amongst existing capacity. This seems to be a reasonable assumption with a base year of 2007. Prior to 2007 there was little variation, which some exception, in capacity growth (i.e. no large difference in rate of return). This assumption may need to be changed in the future with renewables and gas displacing coal and oil power.

The foremost difficulty in validating capacity expansion arises from the lag between planning period to expected service (i.e. lead time) and the changes in economic conditions during that time. Tidball et al. (2010) reports a roughly four year lead time from capacity order to expected service; nuclear and coal may be longer, but do not play a big role in capacity expansion from 2007 to 2018.¹⁰ In reality, there is no natural experiment where an economic shock takes place and the external environment does not evolve while constructing planned capacity. This is especially true in light of the economic recession from 2007–2009 and the decline in gas prices as a result of the US shale oil and gas boom.

This is not to say that a validation is not possible; instead, a validation relies on qualitative discussion of factors leading to the departures between the model outputs and observations. The model results are based on shifts from the 2007 base year, similar to the capacity utilization validation. Nuclear and hydroelectric power are assumed not to expand in the validation, because both are highly constrained by regulation and resource availability, respectively and do not respond as quickly to economic variables. The other technologies are compared to observed capacity from 2007 to 2013 and planned capacity from 2014 to 2018 (EIA, 2015b). We present three validation exercises, each with slightly different assumptions to reinforce confidence in the model results.

The intent of the first validation, termed “targeted total capacity”, is to test how the model performs in predicting contributions from each generating technology by controlling for *total* capacity expansion. Here, capacity expansion in time t is based on the three driving factors: i) technology available during planning at time $t - 4$, ii) assumed perfect information on input prices at time of expected service t , and iii) a projection of generation needs that controls for actual total capacity growth. The first assumption is reasonable because materials must be purchased well in advance of expected service; therefore, the technological efficiency and capital

¹⁰Project lead times vary by technology. Tidball et al. (2010) report an average time from order to expected service of 6 years for nuclear power plants, 4 for coal, 4 for hydroelectric, 2-4 years for gas (depending on technology), 3 for wind, and 2-3 for solar.

costs of the generating units are based on the planning year, time $t - 4$. However, the latter two assumptions may be more contentious, so they are relaxed further in the subsequent two validation exercises: “planning year prices” and “projected capacity needs” validations, respectively.

5.5.1 Targeted Total Capacity Validation

In this validation exercise, the shifts and parameters from Tables 5.3 from the capacity utilization validation and 5.5 below are imposed on the model.

Table 5.5.: Shocks to key drivers of “targeted total capacity” validation from 2007 to 2018. Total generation, \widehat{Q}^g , are exogenously given to target observed total capacity expansion. Policy-adjusted shifts and parameters from Table 5.3 are included in the validation but not shown in this table.

	Exogenous shifts (percentage deviation from 2007)							
Year	Planning ($t - 4$)	Service (t)						
	\widehat{t}_{solarp}^c	\widehat{y}	\widehat{pop}	\widehat{p}_{coal}	\widehat{p}_{gas}	\widehat{p}_{oil}	$\widehat{p}_{O\&M}$	\widehat{Q}^g
2008	30	-0.947	0.950	12.602	22.148	45.969	2.442	0.000
2009	20	-4.639	1.839	20.647	-35.582	-5.395	6.760	2.100
2010	10	-2.728	2.695	21.922	-31.942	26.491	10.082	2.700
2011	-	-1.723	3.483	24.440	-38.820	60.409	13.766	3.600
2012	-10	-0.384	4.276	21.407	-56.569	57.157	14.019	5.200
2013	-20	0.994	5.069	17.643	-45.807	43.594	11.419	1.400
2014	-25	1.396	5.853	17.643	-45.807	43.594	12.382	1.400
2015	-30	2.061	6.852	17.643	-45.807	43.594	13.346	1.000
2016	-33	2.727	7.729	17.643	-45.807	43.594	14.309	1.000
2017	-36	3.392	8.606	17.643	-45.807	43.594	15.272	0.800
2018	-40	4.058	9.482	17.643	-45.807	43.594	16.236	0.600

Figure 5.10 shows that the validation controls for total capacity expansion using the total generation projections, \widehat{Q}^g , shown in the final column of Table 5.5. The

economic recession from 2007–2009 stifled electricity demand. We can observe some of the lagged effects in Figure 5.10. Because the recession was not foreseen and capacity expansion occurs over a 3–5 year time frame, capacity continued to enter service after these years despite the lack of electricity demand and true capacity needs. The forecasts were revised during the recessions causing the capacity to decrease from 2012–2013 before rising again after the recession ended. These confounding factors show the need to control for total capacity growth in this validation. In the broader context, it shows that the electricity sector is highly dependent on linkages in the rest of the economy and reinforces the motivation of this dissertation to understand the role of electricity in the context of the economy as a whole.

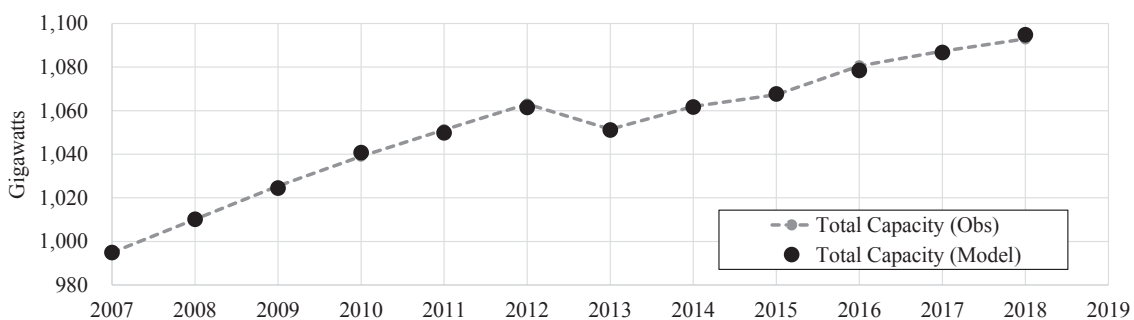
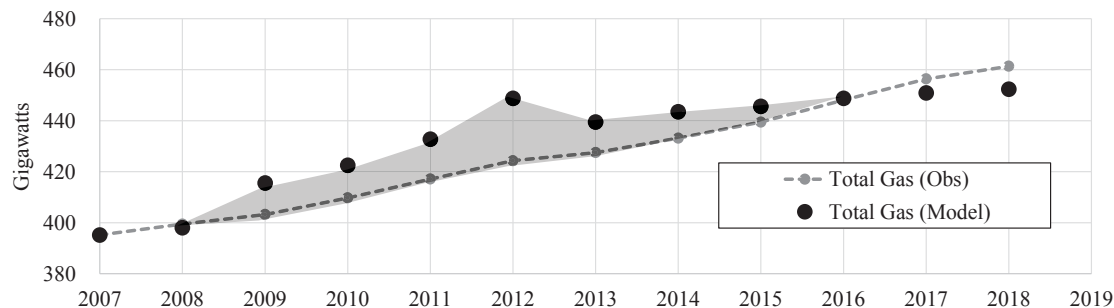
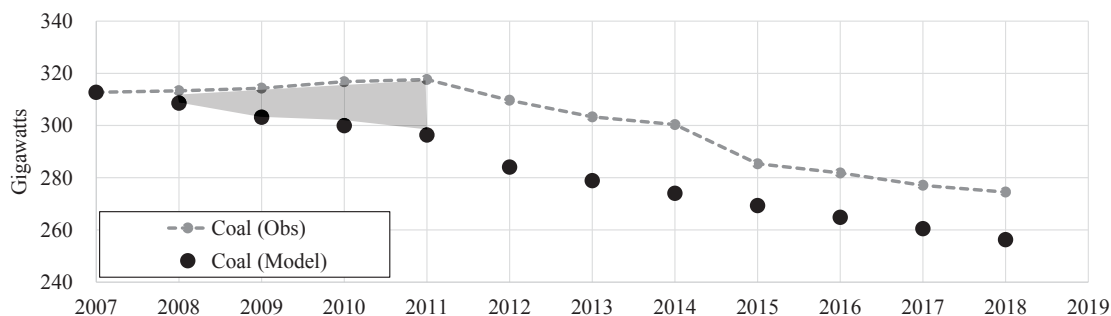


Figure 5.10.: The “targeted total capacity” validation controls for total capacity expansion. Model results are for each year are changes from the 2007 baseline.

As hypothesized earlier, the “targeted total capacity” validation using *service year* input prices over-predicts the observed capacity growth in gas power. This is because the model gives the optimal capacity expansion given *service year* prices, while investment decisions would actually be made in the *planning years* ($\approx t - 4$) when prices of gas were relatively higher (see Figure 5.7). This deviation is represented by the shaded area in Figure 5.11a. Similarly, because gas and coal are highly substitutable (shown in the capacity utilization module) we observe a faster decline in coal capacity – represented by the shaded area in Figure 5.11b. In years after the fall in gas prices the model predictions for gas power and the predictions for the rate of coal retirements more closely mirror those of the observations. This validation fails



(a) Gas capacity expansion in “targeted total capacity” validation. Model results are for each year are changes from the 2007 baseline.



(b) Coal capacity expansion in “targeted total capacity” validation. Model results are for each year are changes from the 2007 baseline.

Figure 5.11.: Gas prices unexpectedly fell beginning in 2008, so due to assumption of input prices at service year instead of input price at planning year, the model over-predicts expansion of gas capacity. The model predicts a more immediate turn from new coal capacity while there is some lag in the observed data.

in a predictable way which lends support for the validity of the capacity expansion in response to fuel prices. The “planned year prices” validation uses *planned year* prices and corrects for some of the deviation in gas power.

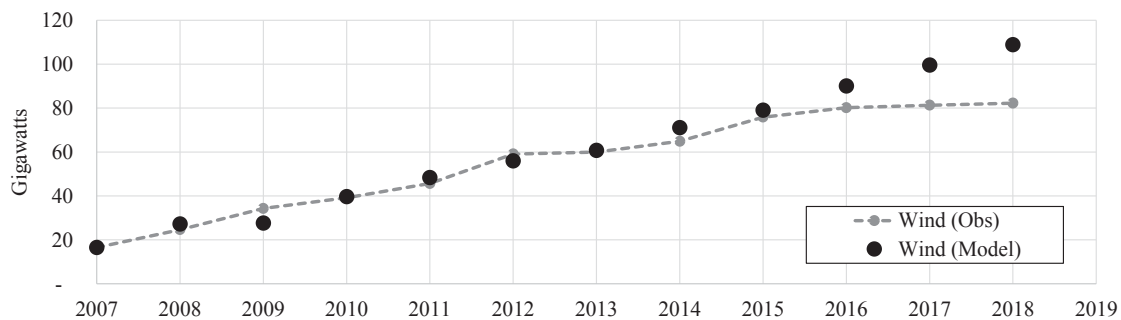
Another point to note regarding coal retirements is that they may not respond immediately to economic stimuli. In this model the base annual rate of retirement is the technical rate of retirement (inverse of technical lifetime). The rate of retirement increases if the returns to capacity decline, and vice versa, to reflect the economic lifetime (see Equation 5.23). In reality, coal retirements may not be so price responsive. The annual rate of retirements and planned retirements observed in the data from 2007 to 2018 is roughly 1.1% of total capacity which implies an

economic lifetime of nearly 90 years, well over their technical lifetime (roughly 60 years), despite decreasing returns to capacity. This may be due to the fact that many of these plants are already paid off and environmental policies preclude the construction of replacement coal plants. There is also uncertainty whether the recent decline in gas prices will continue over the long-term. Coal power operators may elect to produce electricity using existing plants as long as possible, even by co-firing with gas or biomass, to hedge against future fuel price uncertainty and to avoid costly capacity expansion. It may be useful to treat coal power retirements exogenously in long-run analysis, especially in analyzing environmental policy where coal is a significant contributor.

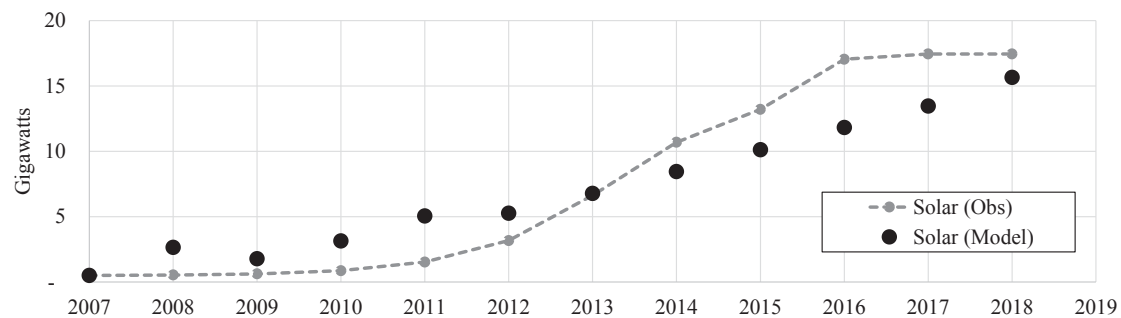
One of the most important trends in the US electricity sector is the expected rise in renewables in response to both technological change and GHG policy. Figure 5.12 shows the model performs fairly well for both technologies despite the rapid growth observed from 2007 onward which is difficult to predict with certainty. Correlations for each technology are shown in Table 5.6.

5.5.2 Planning Year Prices Validation

As we observed in Figure 5.11a using *service year* prices leads to over-predictions of capacity expansion toward gas power. The “planning year prices” validation attempts to correct for some of this by using *planning year* prices ($t - 4$). This might better reflect the initial planning decisions, but neglects the opportunity to adjust capacity expansion plans during the construct phase (e.g. canceling contracts). We might expect reality to lie somewhere in between the two price assumptions. Figure 5.13 shows that using planning year prices in the model provides a better fit to the observed capacity growth. It also shows an outlier in 2012 (i.e. 2008 prices). This is because the model is sensitive to fuel prices. As mentioned, there could have been some year-to-year adjustment in response to reduced gas prices after 2008 (i.e. canceling other contracts and more gas capacity planning in 2009 and after that would have



(a) Wind capacity expansion in “targeted total capacity” validation. Model results are for each year are changes from the 2007 baseline.



(b) Solar capacity expansion in “targeted total capacity” validation. Model results are for each year are changes from the 2007 baseline.

Figure 5.12.: The model does well to predict capacity expansion in renewable power.

come into service in 2012). Generally, equilibrium analysis is useful for long-run planning, so any price shock used in a normal equilibrium modeling scenario would be expected to be a sustained shock and this outlier does not discredit the model’s long-run predictive power.

The “planning year prices” validation in combination with the “targeted total capacity” validation reinforces the assertion that the electricity-detailed model can reasonably predict the contributions to total capacity expansion from individual generating technologies.

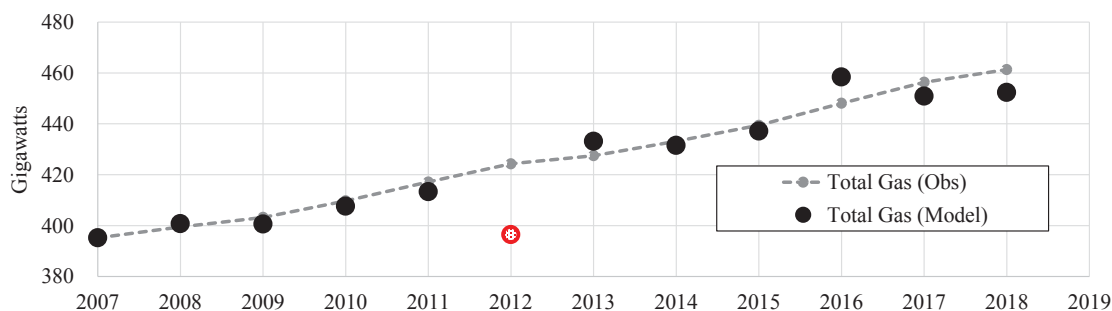


Figure 5.13.: Gas capacity expansion in the “planning year price” validation. Using planning year input prices corrects for the over-prediction using service year prices, but does not allow for year-to-year planning adjustments. Model results for each year are changes from the 2007 baseline.

5.5.3 Projected Capacity Needs Validation

The “targeted total capacity” validation controls for total capacity expansion; however, in a practical modeling scenario the total capacity expansion would be endogenously derived based on the growth in electricity demand. The “projected capacity needs” validation uses a four-year rolling average of generation predictions to endogenously determine total capacity expansion.

The intent of using a four-year average for projected generation is to control for year-to-year variations such as the drop in electricity demand during the economic recession from 2007–2009. Figure 5.14 shows that observed electricity generation was

almost always below the EIA AEO predictions. Viewed in light of the long-term planning horizon in capacity expansion, an over-expansion in the sector would be expected until roughly four years after the recession in 2007–2008; this is observed in Figure 5.10.

Because the AEO projections consistently over-estimated generation needs from 2008–2014 we would expect the model to also over-estimate the actual capacity expansion to some extent. The shaded area labeled “A” in Figure 5.15 shows this to be largely the case. We also observe some deviation for years 2017 and 2018 (the shaded area labeled “B” in Figure 5.15). This is likely due to the fact that planned capacity for 2018 in the EIA data does not account for all the total capacity needs, since it is not necessary to plan four years in advance for some technologies (e.g. wind, solar) which is also why we also observe a plateau in planned expansion of these technologies in 2017 and 2018 as well (Figure 5.12).

The “projected capacity needs” validation shows that the model can reasonably predict total capacity expansion over the long-run, given reasonable projections in total electricity generation.

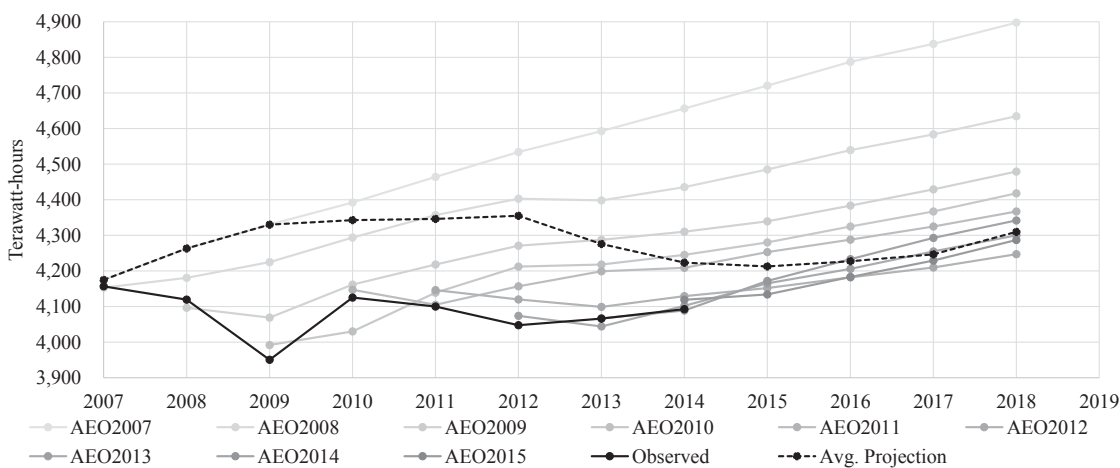


Figure 5.14.: Model results for each year are changes from the 2007 baseline.

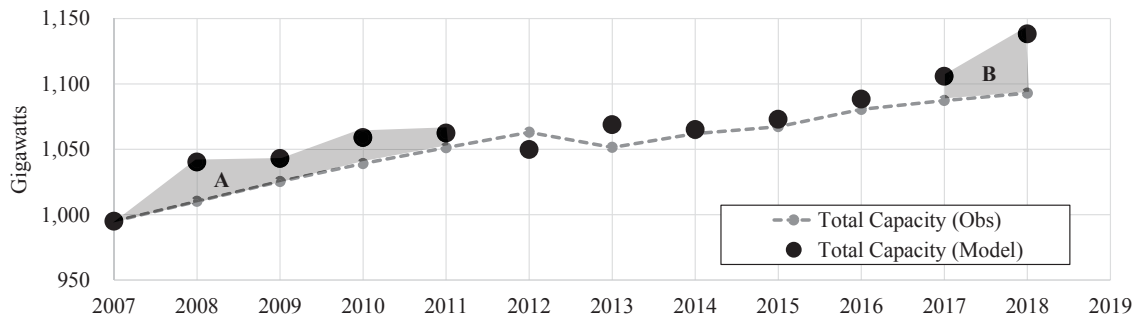


Figure 5.15.: Model results for each year are changes from the 2007 baseline.

5.5.4 Overall Validation

The intent of the three validation exercises is to test the ability of the model to reasonably predict total capacity expansion as well as the individual contribution from individual technologies. Figure 5.6 shows that all of the validations show high levels of correlation between the model predictions and the observed values for contributions from individual technologies with some exceptions. The “other” power sector is poorly correlated. This may be because the technologies which comprise this aggregate group (e.g. geothermal, municipal waste, biogas) are resource constrained, similar to hydroelectric and nuclear, and do not respond as quickly to economic conditions as other technologies.

The first validation exercise, “targeted total capacity” validation showed that the model reasonably predicts the contributions to total capacity expansion for each technology. There were expected deviations for gas power expansion because of the *service year* fuel price assumption, which was relaxed in the subsequent “*planning year price*” validation.

The shale gas boom and subsequent drop in gas price was largely not predicted by the power sector planners. Therefore, the “targeted total capacity” validation that uses prices at time of *service* would be expected to overestimate the initial shift to gas power from coal power. This hypothesis is clearly apparent in the results; that is, the “targeted total capacity” fails in an expected way given additional qualitative information. The “*planning year price*” validation instead used input prices from

Table 5.6.: Comparison of correlation between the targeted total capacity (TTC), planning year prices (PYP), and predicted capacity needs (PCN) validations. Each validation has limitations arising from the assumptions, but each lends support to the overall validity of the model. *The low value in correlation here is due to a single outlier.

	TTC	PYP	PCN
Coal Power	0.916	0.942	0.924
Total Gas	0.898	0.919	0.926
Total Oil	0.978	0.929	0.972
Wind	0.974	0.976	0.901
Solar	0.967	0.626*	0.941
Nuclear	-	-	-
Hydroelectric	-	-	-
Other	0.489	-0.130	0.225
Total Capacity	0.999	0.999	0.908

the *planning year*, $t - 4$. This validation no longer shows the initial overestimate of gas power, but still is not quite reasonable because it does not allow for year-to-year adjustment (e.g. canceling contracts) once price changes are realized during the construction period. The “planned year prices” validation shows correlation gains for gas and coal power which numerically supports the observations in Figures 5.11 and 5.13. Reality is likely somewhere in between these alternate input price assumptions. In a modeling scenario, a price shock would be imposed with the presumption it is a long-term price change; both validations serve to support confidence in price response. Also, the low correlation for solar in the “planning year prices” validation is due to a single outlier discussed in the previous section. The correlation increases to 0.85 when that outlier is removed. This demonstrates the sensitivity of the model and raises some concern regarding fluctuating input prices but is of less concern in long-run analysis.

The “targeted total capacity” validation controls for total capacity so the shares of capacity growth by technology can be seen clearly; however, in a normal modeling scenario the total capacity needs would need to be endogenously derived from some economic shock (e.g projection of generation needs or preferably from projections of income, population, and technical efficiency). Therefore, the final validation, termed the “projected capacity needs” validation, assumes the projected generation for year t to be the average of predictions for year t from the EIA Annual Energy Outlook (AEO) for years $t - 5$ to $t - 2$. The four-year rolling average of projections controls some of the confounding year-to-year adjustment mentioned above and returns total capacity expansion which is shown to correlate well with observations with predictable deviations. The “projected capacity needs” validation shows a high level of correlation between model predictions and observations of total capacity expansion.

5.6 Summary

The electricity-detailed partial equilibrium model presented in this chapter fits well in the space between bottom-up models which do not capture endogenous price feedbacks and economy-wide impacts and top-down models which have insufficient detail in the electricity sector for many relevant technology and policies. The model captures important mechanisms for substitution of electricity generating technologies: capacity utilization, expansion, and their interdependency. Both utilization and expansion are validated against observations of generation and capacity in the United States and perform quite well - especially compared to other economy-wide capable models that are rarely validated in practice. The capacity utilization validation showed strong correlation for generation changes with existing capacity (See Figures 5.8 and 5.9 and Table 5.4). The capacity expansion validation showed an ability to predict contributions to total capacity expansion for individual technologies as well as an ability to capture total capacity changes (see Table 5.6). These lend support for the predictive power of the electricity-detailed partial equilibrium model.

Therefore, we can conclude that the model overcomes two key criticisms of global economic models offered by DeCanio (2003) who argues that: i) models are rarely validated against actual observations, and ii) models do not have sufficient sector-level detail. Despite the demonstrated predictive power, more work may be needed to rigorously estimate the parameters selected for the current version of the model, which will hinge on the availability of longer time series of data.

The following chapter shows the applicability of the model in answering a question for which it is well-suited. What are the different electricity generation mixes resulting from a carbon tax and from a wind and solar investment subsidy used to meet the US Environmental Protection Agency's Clean Power Plan?

CHAPTER 6. ON THE ELECTRICITY SECTOR RESPONSE TO THE CLEAN POWER PLAN: CARBON TAX VERSUS INVESTMENT SUBSIDY

While it has been shown that electricity is an important factor in economic development, the sector outputs potentially harmful greenhouse gas (GHG). An estimated 41%, or 12.5 billion tons, of the worldwide CO₂ emissions from fuel combustion comes from electricity and heat production (IEA, 2012). The US EPA estimated the electric sector was responsible for about one-third of US GHG (CO₂ equivalent) emissions in 2011. In addition to the direct greenhouse gas emissions from electricity production, the life-cycle extends to the upstream (e.g. exploration, transportation, construction) and downstream sectors (e.g. decommissioning, waste management). Mitigating climate change by reducing carbon emissions is on the minds of policymakers worldwide, and the electricity sector is critical to successful mitigation policy.

On August 3, 2015, the US EPA announced the Clean Power Plan (CPP) to reduce carbon pollution. The final rule promotes flexibility in meeting carbon targets by focusing on emission performance that reflects the “best system of emission reduction” based on three building blocks for supply-side management: improved plant-level (namely coal-fired power) efficiency, switching from coal to gas power with existing plants, and constructing more renewable power (EPA, 2015). The second building block pertains to capacity factor adjustment and the third to capacity expansion. Therefore, the partial equilibrium model detailed in Chapter 5 is ideally-suited to study these two interacting mechanisms for meeting requirements of the CPP under the backdrop of the dynamic energy landscape in the United States. Recall that bottom-up models neglect economy-wide impacts and largely consider both fuel prices

and capacity factor to be exogenous while typical top-down models lack the necessary detail to study these dual mechanisms in the electricity sector.

Table 6.1.: Connecting the EPA CPP building blocks to model mechanisms determining changes in electricity generation (EPA, 2015).

	Description from EPA (2015)	Model Mechanism
Building Block 1	reducing the carbon intensity of electricity generation by improving the heat rate of existing coal-fired power plants	Technical productivity (\hat{a}_t^g)
Building Block 2	substituting increased electricity generation from lower-emitting existing natural gas plants for reduced generation from higher-emitting coal-fired power plants	Capacity utilization (\hat{c}_t)
Building Block 3	substituting increased electricity generation from new zero-emitting renewable energy sources (like wind and solar) for reduced generation from existing coal-fired power plants	Capacity expansion (\hat{q}_t^c)

Previous drafts of the Clean Power Plan rule included a fourth building block: more efficient electricity use. Increasing end-use efficiency may have significant potential in offsetting total electricity demand in the United States (Wang and Brown, 2014). The focus of this work is on endogenous supply-side responses (i.e. capacity utilization and expansion); both supply and demand efficiency are considered implicitly as the difference between gross national income and an exogenous projection of electricity demand (Equation 5.19).

The intent of the rule is to give individual states flexibility to meet their state-specific targets. The basic options are to establish a carbon trading system within or across state boundaries (i.e. addressing building blocks 2 and 3) or establish subsidies or regulations for new capacity (i.e. addressing building block 3 only). While the CPP is aimed at individual states, the emission target is at the national-level. The CPP, when fully implemented, aims to reduce carbon pollution from the US electricity sector by 32% below 2005 levels.

In order to give some perspective of the CPP target, it is necessary to observe some important trends in the US electricity sector, especially in terms of some changes in CO₂ intensity, presumably due in part by the same mechanism discussed in the CPP building blocks. Figure 6.1 shows that CO₂ intensity remained fairly constant from 1989–2007, and decreases significantly after 2007 (see area A in Figure 6.1). This could be a result of fuel-switching from coal to gas power (coal being almost twice as carbon intensive as gas) as observed in Figure 5.9, but could also be a result of the increase in renewable capacity as observed in Figure 5.12 or several other factors such as general increasing carbon efficiency. If we use the total generation projection from the EIA Annual Energy Outlook (AEO) 2015 report (EIA, 2015a) combined with electricity sector-wide projections for CO₂ intensity using data from 1989–2013, we observe a reduction of 8.2% in total CO₂ emissions from the 2005 baseline. If we instead use a projection of CO₂ intensity using data from just 2007–2013, where we observe a decoupling in generation and emissions, the reduction of total CO₂ emissions are over 44% from the 2005 baseline. While reality likely lies somewhere in the middle, Figure 6.1 shows that the momentum from current economic conditions (e.g. low gas prices, renewable policies, and end-use efficiency) will likely reduce the magnitude of policy interventions needed to meet the CPP target.

The analysis begins by establishing the baseline for 2030 based on current input prices, policies, and projections of future generation requirements. That is, given the current US energy-economic landscape with low gas prices, renewable policies, and increasing electricity end-use efficiency, how close to (or how far from) the 32% target will we end up in 2030. What we find is that from increasing returns to gas capacity due to the increasing capacity utilization observed in Chapter 5 and the opposite trend in coal, emissions are projected to decrease almost 28% from 2007 levels without additional policy instruments.

Still, policy instruments may be needed to make up the full CPP CO₂ target. Sections 6.2 and 6.3 contrast the impacts of a carbon tax (equivalent to carbon trading mechanism) with an investment subsidy for only wind and solar power (W+S

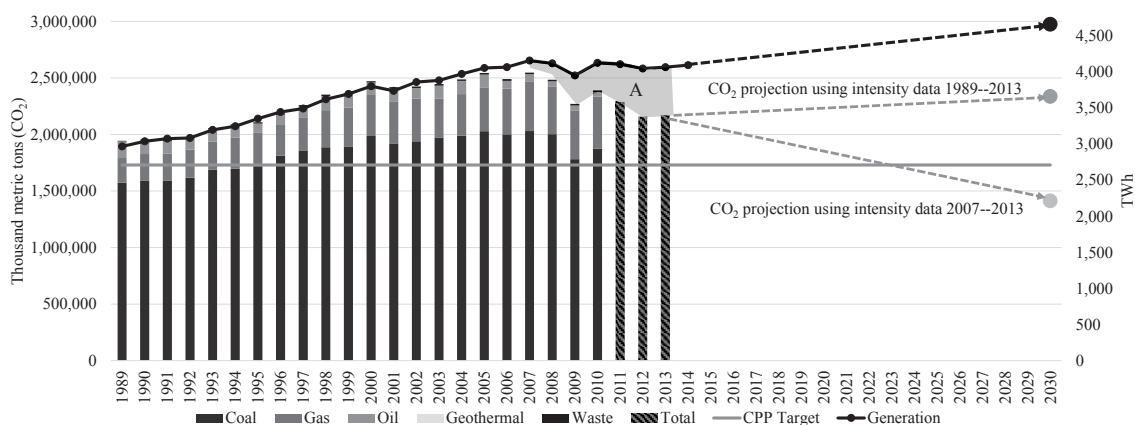


Figure 6.1.: CO₂ emissions by source for the US electricity sector from 1989 to 2010, and as total emissions from 2011 to 2013 (EIA, 2015b). We observe decreasing CO₂ intensity from 2007 onward (A). The total generation projection for 2030 is based on EIA AEO 2015 (EIA, 2015a), and the two total CO₂ projections are based on CO₂ intensity projections using linear regressions with different periods of the emission data (1989–2013 and 2007–2013).

investment subsidy scenario) to meet carbon reduction targets, and also compares the resulting generation mixes in each scenario. Section 6.4 concludes.

6.1 Baseline for 2030

The baseline scenario is used to provide a basis for analyzing the carbon tax and the W+S investment subsidy scenarios. The scenario projects the electric power sector to 2030 using projections for total generation needs using the most recent EIA AEO (EIA, 2015a), as well as population (US Census), labor costs as a proxy for O&M costs (BLS, 2015), and income per household using simple regressions. With the exception of large shocks such as the US economic recession discussed in Chapter 5, growth rates for these projections are quite stable. We shift fuel prices, taxes, and technology to 2014 levels. The fuel prices still have a supply response. Given recent fluctuations in fuel prices resulting from the shale oil and gas boom and the more recent decline in oil prices (2015), no change may be a satisfactory projection. Further, empirical evidence on endogenous technological change suggests that the

price of fuels remains fairly constant in real terms over the long-term, albeit with high year-to-year variability (Stuermer and Schwerhoff, 2013). Current federal investment subsidies for wind and solar (i.e. 30% capital subsidy) are included in the baseline simulation. Because technology is assumed to remain at 2014 levels, the results may be conservative estimates for relatively new technologies (e.g. wind and solar) that are improving at a faster rate than traditional technologies. These exogenous shifts to the 2007 base year are shown in Table 6.2.

Table 6.2.: Shocks to 2030 for baseline scenario based on projections or 2014 observations. Effective taxes from mercury regulation (Table 5.3) and technology (Table 5.5) are also shifted using 2014 levels. Years 2014 (fuel price year) and 2018 (last year of joint validation) used for reference.

	Exogenous shifts (percentage change from 2007)							
	Projections				2007–2014 change			Model Output
Year	\hat{y}	\widehat{pop}	\hat{Q}^g	$\hat{p}_{O\&M}$	\hat{p}_{coal}	\hat{p}_{gas}	\hat{p}_{oil}	\hat{a}^g
2007	-	-	-	-	-	-	-	-
2014	1.396	5.853	1.400	12.382	17.643	-45.807	43.594	5.360
2018	4.058	9.482	0.600	16.236	17.643	-45.807	43.594	9.015
2030	12.044	20.004	12.853	27.765	17.643	-45.807	43.594	17.423

Furthermore, coal and oil capacity are assumed to not expand economically due to regulatory constraints (e.g. mercury and carbon regulations). Nuclear and hydroelectric capacity are assumed to not expand economically due to regulatory and resource constraints, respectively.

There will be additional retirements which will need to be replaced with additional capacity. Looking back at the results we observed the joint validation in Chapter 5 and because we are using the same fuel price shifts, it is reasonable to assume that most of the expansion in the 2030 baseline will also come from the gas, solar, and wind technologies. We can also hypothesize that because capacity utilization changes in 2030 should roughly match those from the projected capacity validation for 2018

due to identical fuel price shifts, returns to capacity should be roughly similar and expansion would probably be in roughly similar *shares*, despite the highly non-linear system of equations. This hypothesis does not consider confounding factors from the capacity expansion.

The results shown in Figure 6.3 and 6.4 support the above hypothesis. We observe that capacity utilization of ‘GasBL’ increases over 60% while coal power utilization decreases due to the lower relative price of gas. The story is similar for the reduced utilization for oil power. Interestingly, we also observe a slight decrease in utilization of ‘GasP’ which is likely a result of the tremendous growth in solar.

Utilization, in turn, affects the returns to capacity. Pressure from increasing utilization and capacity expansion in ‘GasBL’ and expansion of wind puts pressure on inflexible technologies that cannot adjust economically (i.e. nuclear and hydro). Significant retirements of both coal and oil capacity (decreasing supply) keeps returns fairly high for remaining plants, despite their lower utilization rates and inability to expand.

The interdependency of utilization and expansion are nicely demonstrated for ‘GasBL’ in Figures 6.3 and 6.4. Returns to ‘GasBL’ fall moving out to 2030 due to significant expansion in capacity. With high returns in 2018, ‘GasBL’ encompasses 23% of total capacity expansion, but due to the large expansion some of the returns are “crowded-out” and the share reduces to 15% in 2030.

As we observe in the validation exercises in Chapter 5, there is strong switching from coal power to gas power. Also, additional coal plants retire with time. As a result, total carbon emissions from the US electricity can be expected to drop approximately 27.7%. Total carbon emissions from coal and oil power reduce by 48% and 60%, respectively, but is partly offset by a 72% increase in emissions from gas power. Figure 6.2 shows carbon emissions by source in 2030 compared to 2005.

Interestingly, the baseline scenario projects that with no additional policies the CPP goals will almost be met by 2030 (27.7% of the 32.0% target). Only a

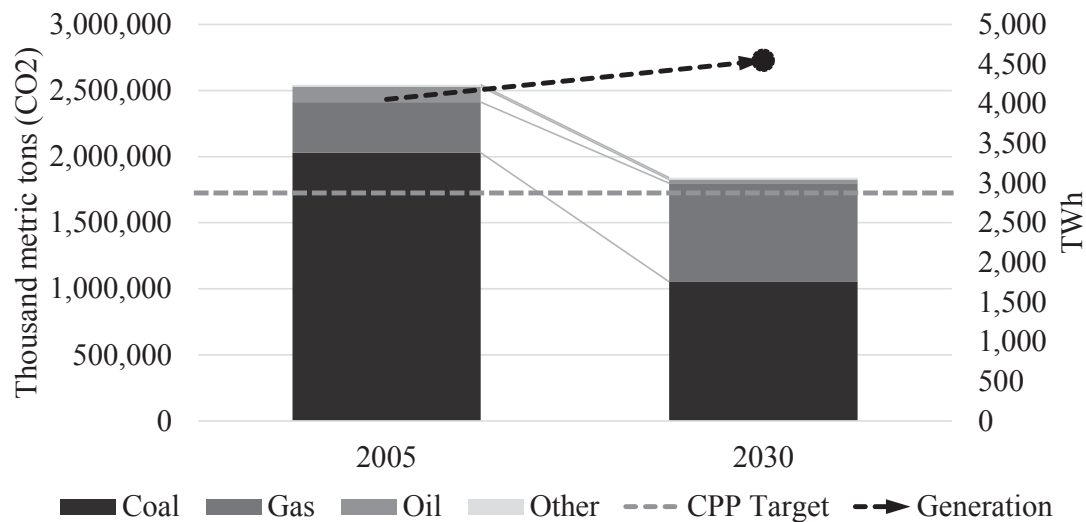
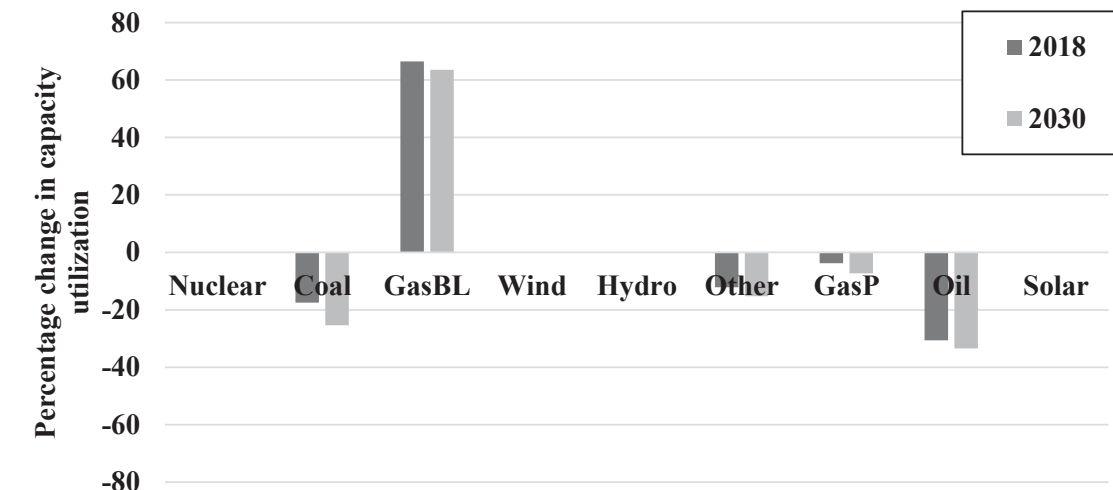
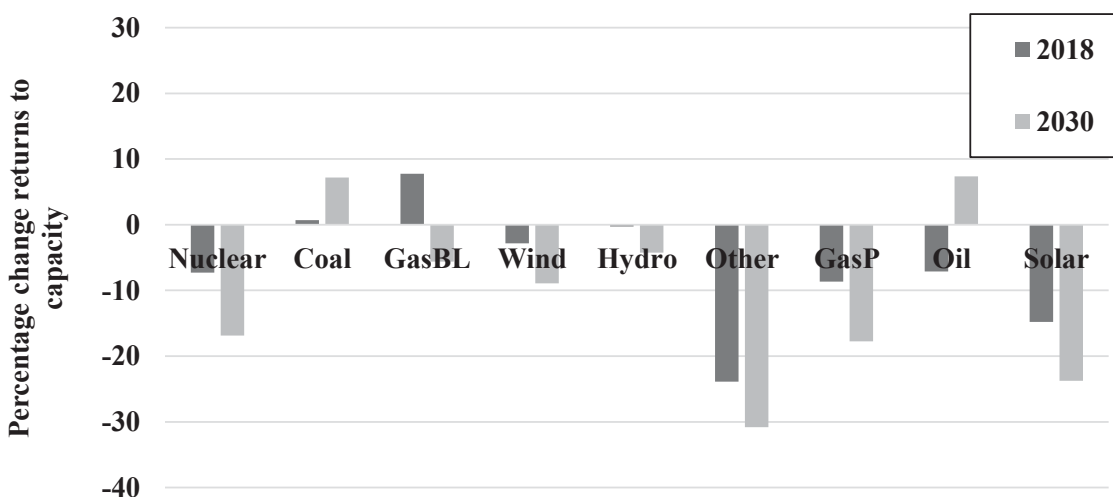


Figure 6.2.: Contributions to total CO₂ emissions by fuel-type in the United States in 2005 and model projections for 2030. Emissions from coal and oil reduce by 48% and 60%, respectively, but is partly offset by a 72% increase in emissions from gas power. CO₂ emissions for other power technologies is small.



(a) Capacity utilization changes from 2007–2018 compared to 2007–2030



(b) Returns to capacity changes from 2007–2018 compared to 2007–2030

Figure 6.3.: Capacity utilization in 2018 and 2030 are roughly similar due to similar fuel price shifts (dissimilar O&M price shifts). This results in returns to capacity in 2018 and 2030 are roughly similar except for technologies that have difficulty expanding economically due to policy (i.e. coal and oil) or resource and regulatory constraints (i.e. nuclear and hydro).

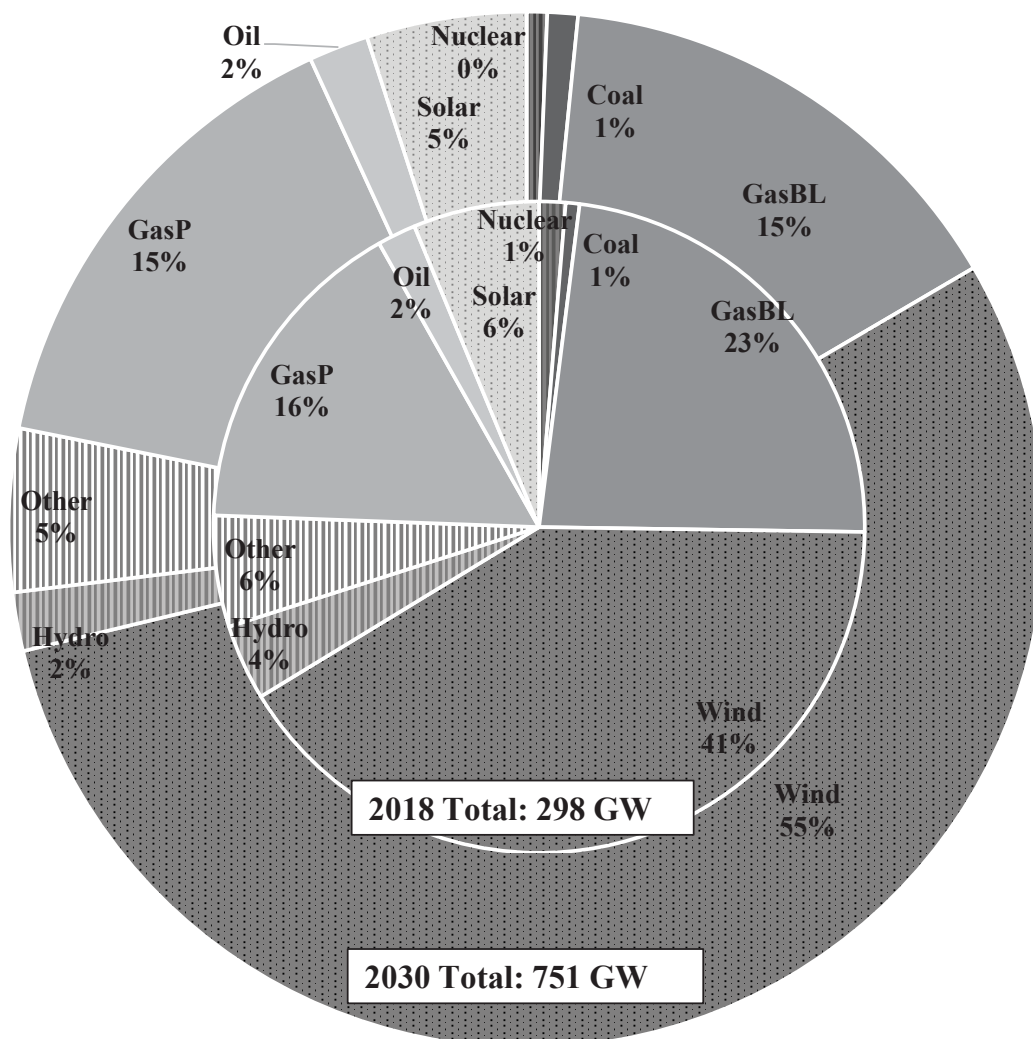


Figure 6.4.: The shares of **additional** nameplate capacity for 2007–2018 and for 2007–2030 are roughly similar. Areas to scale.

relatively small remainder must be achieved through further incentives through policy interventions which are discussed in the following section.

The 27.7% reduction predicted here is significantly different than other “business as usual” (BAU) cases in both bottom-up and top-down literature. For example, Bushnell et al. (2014) use a bottom-up, partial equilibrium model to analyze differences between cap-and-trade and rate-based policies in a handful of western US states. Their BAU case estimates a 10% increase in emissions from 2007–2030 due to increasing demand because the partial equilibrium model does not consider increasing returns to capacity due to utilization changes, which would drive expansion toward gas and away from coal, as in the model described here. This interdependency between utilization and expansion is usually ignored in least cost optimization models. As a result the policy intervention needed to meet the CPP target is larger than what is found in the following chapter. On the other hand, Cai and Arora (2015) use a CGE model with several generating technologies which substitute imperfectly, which implicitly captures a combined utilization and expansion effect. The 2030 baseline in this study predicts an 11% decrease in CO₂ emissions. However, characteristic of most top-down models, the model does not consider important aspects of electricity production (e.g. base versus peak markets), is not validated, and includes technologies that do not currently exist in meaningful scale (e.g. carbon capture and storage).

Without considering the interdependency of utilization and expansion (as in many bottom-up optimization models) estimates of emissions may be too high. Without explicitly representing and validating these joint mechanism the numbers resulting from top-down models should be met with appropriate caution. The representation and validation of capacity utilization and expansion in this model lends credibility to the estimate of CO₂ emissions presented here and in subsequent sections.

6.2 Policy Scenarios: Carbon Tax and W+S Investment Subsidy

The following policy scenarios target the full 32.0% reduction goal of the CPP in the US electric power sector. First, a carbon tax on electricity generation technologies, T^{CO_2} , of \$8.9 per metric ton of CO₂ is imposed in order to reach the policy target. Second, equivalent capital (i.e. capacity) subsidies, \hat{t}_t^c , of -27 are given to wind and solar exclusively. The two policies are equivalent in that they are required meet the additional 4% reduction needed to meet the full 32.0% reduction in CO₂ emissions from 2005 levels in 2030. These two alternative shifts are shown in Table 6.3 where $\widehat{CO_2}$ is the percentage change in CO₂ emissions relative to the 2007 economy.

Table 6.3.: Additional shifts to 2030 in addition to baseline shifts in Table 6.2 for carbon tax and W+S investment subsidy policy scenarios. Both policy scenarios choose shocks that meet the total CO₂ reduction goal of the CPP, 32.0%.

	Projections		Policies			Target
Policy	Year	\hat{a}^g	T^{CO_2}	\hat{t}_{solar}^c	\hat{t}_{wind}^c	$\widehat{CO_2}$
Carbon tax	2030	17.423	\$8.9	-	-	-32.0
W+S investment subsidy	2030	17.423	-	-27	-27	-32.0

The two alternative policy scenarios would be expected to have different impacts on the economy due to targeting of specific technologies as well as how the shocks impact the systems of equations which define the model economy.

First, the carbon tax targets CO₂ directly by internalizing the cost of emissions via the zero-profit condition for the electricity generating technologies:

$$\hat{p}_t^g = \sum_i c_{it} \cdot \hat{p}_{it} + \gamma_t \cdot T^{CO_2} \quad (6.1)$$

where the first term is the original zero-profit condition in Equation 5.18 and the second accounts for the carbon tax. The coefficient, γ_t , is the CO₂ emission factor (thousand metric tons per GWh) for technology t . Table 6.4 shows that emission factors are non-zero for coal, gas, oil, and other power, and vary in intensity. The

carbon tax negatively impacts these technologies while positively impacting other technologies on a relative basis. On the other hand, the W+S investment subsidy positively impacts wind and solar while negatively impacting all other technologies on a relative basis. Under both scenarios, fossil fuel technologies are hurt. The difference is in non-fossil fuel technologies that do not receive subsidies (i.e. nuclear and hydro power), which are impacted positively in the carbon tax but negatively in the W+S investment subsidy scenario. This is because zero-emitting technologies, like nuclear and hydro power, are relatively better options under a carbon tax, but are not included in the wind and solar only subsidies so they are hurt in this case. Fossil fuel technologies that are less impacted by the carbon tax than others (i.e. gas and other power) are impacted differentially in the carbon tax case and uniformly (same basis) in the W+S investment subsidy case. Therefore, the less carbon intensive technologies are hurt more in the investment subsidy case.

Table 6.4.: CO₂ emissions factors for each technology in the United States. EIA data for 2007.

	Nuclear	Coal	GasBL	Wind	Hydro	Other	GasP	Oil	Solar
γ_t	0	1.007	0.477	0	0	0.198	0.477	1.031	0

The second difference between the two policy scenarios is how the policies enter the system of equations that define the model economy. The primary linkages are outlined below:

- The carbon tax, T^{CO_2} , will impact the cost of electricity generation via Equation 6.1.
- The costs of generation impact capacity utilization via Equations 5.17 and 5.16.
- Capacity utilization contributes to the returns to capital via Equation 5.4.
- Returns to capacity and \hat{t}_t^c contribute to additional capacity growth Equation 5.28 and Equation 5.29.

- Total capacity changes feed back to the returns to capacity (Equation 5.7) and cost of generation.

The carbon tax will mainly impact capacity utilization which impacts the returns to capacity. The investment subsidies for wind and solar will impact capacity expansion and affect utilization by crowding-out other technologies. The two different policies reinforces the need to capture the dual mechanisms for substitution in generation: utilization, expansion, and their interdependency.

The following section follows the primary linkages listed above to interpret the equilibrium electricity mix in 2030 in response to the two policy scenarios.

6.3 Policy Scenario Discussion

Figure 6.5 shows that the carbon tax increases fuel switching from coal to gas power as compared to the baseline. The relative impact on utilization is correlated with the emission factor – more emissions per GWh more negative impact. Interestingly, utilization of the technologies in the other technology category actually switches from negative in the baseline to positive under the carbon tax because it is relatively less affected than fossil fuel technologies. As expected, the investment subsidies for wind and solar crowd-out capacity utilization for all technologies that are not subsidized in a more uniform manner than the carbon tax scenario. As a result, fossil fuel technologies with lower emission factors (e.g. gas and ‘other’ power) are hurt relatively more than in the carbon tax scenario, fossil fuel technologies with higher emission factors (e.g. coal and oil power) are hurt relatively less (see Figure 6.5).

The story in utilization translates easily to the returns to capacity shown in Figure 6.6. The carbon tax drives coal capacity returns down while the W+S investment subsidy drives wind capacity returns up. The high relative returns for wind crowds-out capacity expansions in gas power which drives returns for gas a little higher, but also crowds-out capacity expansion in solar power. Without

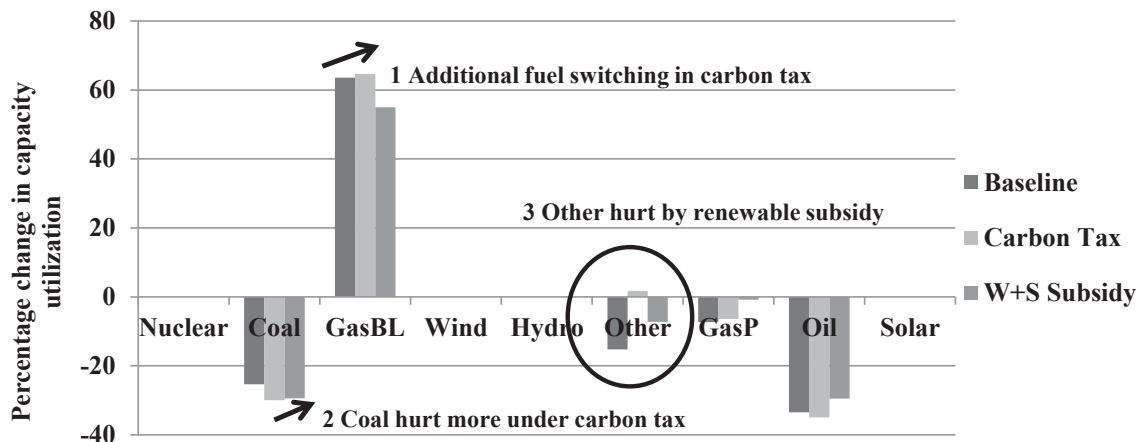


Figure 6.5.: A carbon tax reduces capacity utilization for fossil fuels from the baseline based on relative carbon content (1, 2) while a wind and solar investment subsidy impacts utilization for all other flexible technologies in a more uniform manner (1, 2, 3).

as much expansion in solar, ‘GasP’ remains an important part of the generation mix for peak load, and we observe higher returns in the W+S investment subsidy scenario. Inflexible, non-fossil fuel technologies (i.e. nuclear and hydro power) are impacted positively in the carbon tax case, but hurt tremendously in the investment subsidy case. This is a known phenomena where the policy essentially picks the winning technologies, here mainly wind, and implicitly picks losers despite the losing technologies’ ability to address the stated goal (i.e. reduction in CO₂ emissions).

The resulting generation mix from the two policy scenarios differ in three main ways. First, in the carbon tax scenario all non-fossil fuel technologies expand. This includes nuclear and hydro which would also be expected to expand even more if regulatory restrictions were reduced for nuclear or hydro resources made available. Picking the winners (wind and solar) constrains the returns for these other technologies. Recall, that the model limits movement in these technologies due to their imperfect response to economic stimuli; still, the direction of change is important especially when analyzing the results with this type of qualitative information (as shown in the validation exercises in Chapter 5). Second, the differential taxation

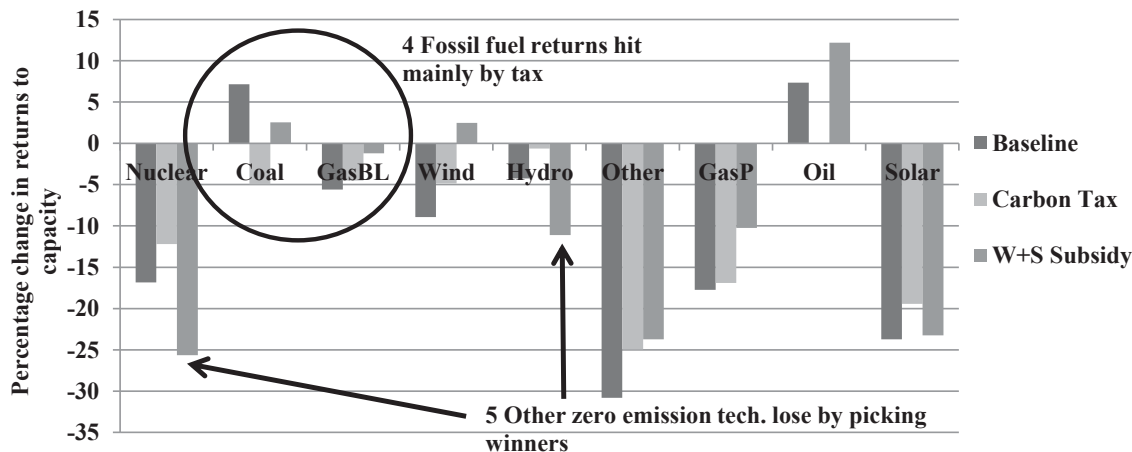


Figure 6.6.: Absolute returns for fossil fuels are reduced by the carbon tax and less impacted by the wind and solar investment subsidies (4). Returns to other zero emissions technologies that not subsidized are hurt by the investment subsidy scenario (5).

of carbon emissions between gas and coal leads to additional fuel switching from coal to gas power; however, there is a decrease in both gas and coal power in the W+S investment subsidy scenario. Third, as expected wind expands more in the subsidy case. Finally, there is significant interdependence in capacity expansion and utilization, as we observe with increasing utilization of ‘GasP’ (because of crowding-out of expansion due to wind capacity returns) despite subsidies for solar power expansion which also leads to a reduction in total solar power.

Recall that the CPP focuses on emission performance that reflects the “best system of emission reduction” based on three building blocks for supply-side management: 1) improved plant-level (namely coal-fired power) efficiency, 2) switching from coal to gas power with existing plants, and 3) constructing more renewable power. Both policy scenarios address the third building block. Renewable (i.e. nuclear, hydro, wind, and solar) capacity increases 5.0% with a carbon tax and 7.9% with W+S investment subsidies. However, the story is slightly different for building block 2 – switching from coal to gas power with existing plants. The baseline case with no additional CO₂ emission policies shows that 41.39% of total

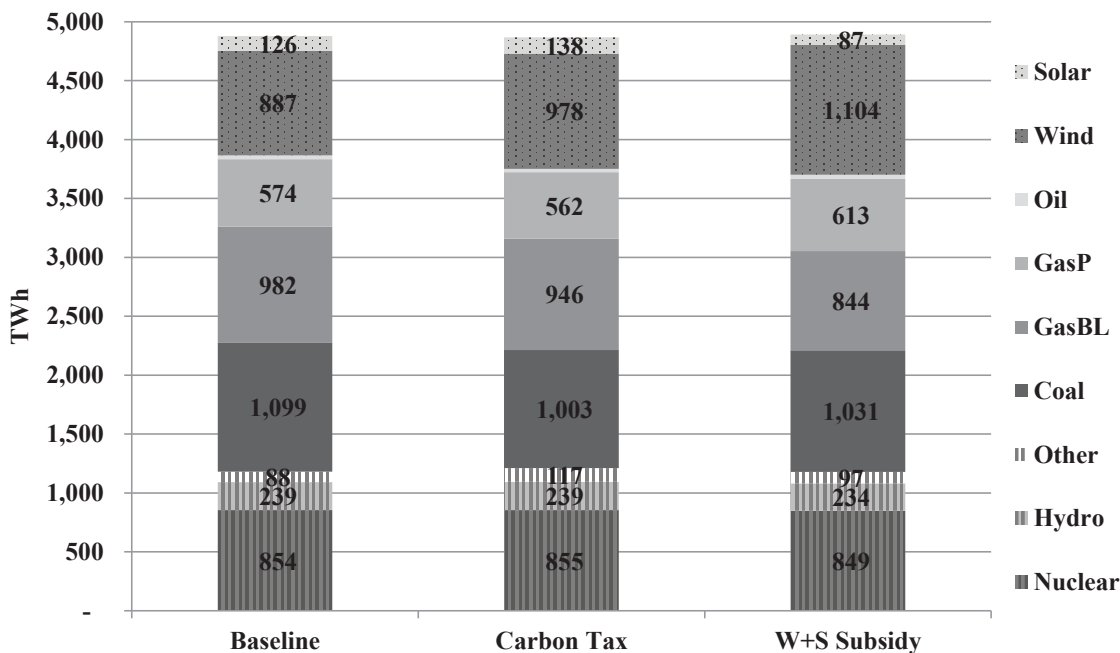


Figure 6.7.: Power generation in 2030. A carbon tax allows some zero emission technologies (i.e. boxes with vertical lines) to expand slightly while these are reduced in the W+S subsidy case. The carbon tax allows for greater switching from gas to coal (CPP building block 2) as compared to the W+S subsidy case. As expected larger growth in wind and solar is observed with the subsidies.

electricity production from either coal and gas power comes from coal power. Coal power's share reduces to 39.94% with the carbon tax which indicates additional fuel switching from the baseline. However, with subsidies to wind and solar the share of coal power increases to 41.45% which indicates that expanding renewable capacity (CPP building block 3) may compromise building block 2.

The cost of total electricity production increases 8.8% in the baseline scenario, primarily due to increasing labor costs. The carbon tax increases the electricity costs an additional 3.0%, which reduces total electricity demand, while the W+S investment subsidy decreases electricity costs by 0.6%, increasing demand slightly (see Figure 6.7). The total electricity cost may have important impacts in both tax receipts and welfare analysis, which are not considered here.

Another important point in this analysis is that the biggest contributor to emissions in the US electricity sector comes from coal power; however, none of the CPP building blocks explicitly call for retiring coal plants. Reducing electricity generation from coal requires reducing utilization as well as reducing capacity. Building block 2 helps drive retirements because differential relative prices (e.g. inexpensive gas) and taxes (e.g. carbon tax) can drive down utilization of coal which also reduces returns to capacity and increases retirements. Because the W+S investment subsidy does not reduce the returns to coal power as much as the carbon tax in absolute terms (Figure 6.6), the rate of coal retirements is lower (50.6% reduction in coal capacity with a carbon tax and only 48.0% with the W+S investment subsidies). As Chapter 5 notes, coal retirements may not be so price responsive, and it may be worthwhile to treat coal retirements exogenously and perform sensitivity analysis on the rates of retirement. In the absence of a carbon tax, subsidies might be more effective if they are designed to provide an additional incentive to retire coal plants to ensure that both building blocks 2 and 3 are met simultaneously.

6.4 Conclusions and Future Work

This chapter applies the partial equilibrium model described in Chapter 5 using the data for the US electricity sector constructed in Chapters 2–4. A carbon tax and an investment subsidy scenario for solar and wind are compared to the baseline. The results replicate well-known results regarding the efficiency of a tax versus selective regulation (i.e. wind and solar investment subsidies, here).

One of the more interesting insights from this work is the baseline projection to 2030. Given current economic conditions, namely the assumption that gas price will remain at 2014 levels in the long-run, almost 28% of the CPP target of 32% reductions will be met without any additional policy intervention. This means the magnitude of further policies (e.g. carbon tax and investment subsidies analyzed here) may be much smaller than what we might expect. Additional research on the “break-even”

cost of gas for meeting the CPP target as well as investigating economy-wide factors that might affect the price of gas such as the opening of LNG exports Levi (2012) would be both timely and interesting. Increased demand for gas by the electricity sector only increased the price of gas about 4.6%.

By capturing utilization, expansion, and their interdependency explicitly in the model we can observe some additional interesting phenomena. First, selective investment subsidies hurt returns for other renewable generating technologies (e.g. nuclear, hydroelectric power) that otherwise could contribute to the goal of emission reductions. Selective piecemeal subsidies can also compromise the effectiveness of policies. Recall, that wind subsidies crowded-out investment in solar, which actually resulted in more fossil-fuel production in the peak load. Second, focusing on the third CPP building block (i.e. capacity expansion for renewables) might actually hurt the second building block (i.e. fuel switching from coal to gas power with existing plants). When pursuing a “best system for emission reduction”, as outlined in the EPA CPP guidelines, it is important to combine capacity expansion policies with other policies which might prevent emission offsetting in fuel-switching. Third, perhaps the best way to reduce emissions is by retiring coal capacity. This is done by reducing the returns to capacity. The carbon tax reduces returns by encouraging fuel switching to gas, but the investment subsidies only reduce coal power returns in relative terms, which does not push coal power out of the electricity mix at as quickly of a rate.

This analysis gives guidance for designing a “best system of emission reductions” to meet the CO₂ emission reduction objectives of the CPP. However, policymakers should also consider welfare and economy-wide effects. Further, CO₂ emissions have implications beyond the just the US region via our shared global economy and environment. Carbon leakage has been shown to be an important consideration for reducing global emissions (Peters and Hertwich, 2008). Also, possible opportunities for fuel exports (i.e. oil, gas, and coal) could open up the electricity sector to important international price dynamics. The next steps in this line of research are

to integrate the computation model as part of a global CGE model with relevant sectoral and regional linkages.

CHAPTER 7. CONCLUSIONS AND FUTURE WORK

The electricity sector plays a crucial role in the global economy. It is also a major consumer of fossil fuel resources, a large producer of greenhouse gas emissions, and an important indicator and causal factor for economic development. As such, the sector is a primary target for policymakers seeking to address each of these issues. The sector is also experiencing rapid technological change in generation (e.g. renewables), primary inputs (e.g. horizontal drilling and hydraulic fracturing), and end-use efficiency. This dissertation seeks to further our understanding of the role of the electricity sector as part of the dynamic global energy-economy, which requires significant research advances in both database construction and modeling techniques.

The introductory chapter discussed how the advances presented in this dissertation fit within the current field research that focuses on the electricity sector and its economy-wide linkages. So called “bottom-up” partial equilibrium or simulation models are capable of providing a technologically and operationally-detailed representation of the electricity sector. However, the sector-level detail often sacrifices inter-industry feedbacks (i.e. fuel and other input prices are considered largely exogenous) and inter-regional linkages. “Top-down” models, such as computable general equilibrium, suffer from the opposite affliction. That is, inter-industry and inter-regional linkages are well represented at the expense of sector-level details. The advances in database construction and representing electricity presented in this dissertation lend capability and credibility to sector-specific modeling in a top-down computational equilibrium framework.

Chapters 2–4 detail individual, but interrelated advances in CGE database construction for global economic modeling. Chapter 2 described the need to balance bottom-up, engineering-level data with top-down data used for global economic

analysis. Previous research that constructed electricity-detailed CGE databases fail to document matrix balancing methods which balance these disparate data. The chapter revisited the well-studied matrix balancing problem and formulates a share preserving cross entropy (SPCE) method that is designed specifically for the type of data in the electricity disaggregation. The solution to this method is equivalent to the well-studied RAS method, but allows one of the constraints to be relaxed without compromising the solution. This means the SPCE method is more flexible than RAS and is shown to perform better in the case where total costs of sub-sectors is not known with absolute certainty, as in the case of the electricity data. Chapter 3 demonstrated that matrix balancing methods have a significant impact on modeling results and advocates for more documentation of these methods across researchers in the future. GTAP-Power leverages these and other advances in constructing the most electricity-detailed publicly-available CGE database (Chapter 4).

Chapter 5 described an electricity-detailed model that leverages the GTAP-Power database to address two of the most criticized aspects of computational equilibrium modeling: i) insufficient sector-level detail and ii) lack of validation against observed data. The model captures two aspects of electricity generation which is the relevant metric for balancing supply and demand as well as determining externalities such as GHG emissions. Capacity factor utilization, expansion, and their interdependency are explicitly and endogenously determined, which advance the field of both bottom-up and top-down computational equilibrium model. The capacity utilization is validated for the United States for the years 2002 and 2012 and shown to perform quite well, especially in combination with qualitative information. A joint validation, which links utilization with expansion, encompasses several distinct simulations which test the models predictive ability from 2007 to 2018 and builds confidence of the system of equilibrium equations in modeling the US electricity sector. Thus, this chapter addresses two fundamental criticisms of computational equilibrium modeling.

Chapter 6 used the GTAP-Power database for the United States and the validated electricity-detailed model to analyze the US EPA Clean Power Plan. The 2030

baseline scenario almost completely achieves the 32% CO₂ emission reduction goal of the CPP. Low gas prices, coal power retirements, and expansion in renewables nets reductions of 27.7% *with no additional policy interventions*. A carbon tax and a solar and wind investment subsidy are imposed on the baseline scenario to reach the full 32% target of the CPP. A wind and solar investment subsidy picks winning technologies at the expense of over renewable technologies. Interestingly, the results show that the wind investment subsidy crowd-out expansion for solar capacity even with the solar investment subsidy. Further, the investment subsidies can offset some of the potential baseline CO₂ emission reductions from switching from coal to gas power, which is one of the specific mechanisms, termed “building blocks” outlined in the CPP. Finally, we observe that the carbon tax penalizes coal power in absolute terms rather than in just relative terms by increasing returns of wind and solar. As a result, coal capacity retires at a slower rate in the wind and solar investment subsidy scenario.

Because coal is emissions intensive, retiring coal capacity is the key component in emission reduction. An investment subsidy might be combined with another mechanism to reduce returns to coal capacity to increase the rate of retirements. In the absence of an economically efficient carbon tax, the more tractable policy of renewable investment subsidies should be designed in such a way to avoid “picking winners” (e.g. subsidies for *all* zero and low-emitting technologies), offsetting fuel switching possibilities, and increasing returns to coal capacity.

Overall, the computational equilibrium model, which captures capacity utilization, expansion, and their interdependency, is shown to be well-suited to analyze EPA’s advice to pursue a “best *system* of emission reductions” given the complex nature of the electricity system.

This dissertation focuses on advances in database construction and representing electricity in economy-wide analysis. The next steps involve integrating the computational equilibrium model from Chapter 5 into a full CGE framework to fully leverage the inter-industry and inter-regional capabilities. The CGE linkages

will expand analysis to income and welfare impacts to more comprehensively analyze carbon mitigation scenarios such as those proposed in Chapter 6. Further, trade linkages will likely become of increasing importance for the US electricity sector as traditional generating fuels (i.e. coal, gas, and oil) have potential for exports. There are currently a handful of LNG export terminal permits approved for exporting gas. There is legislation looking to remove the ban on petroleum exports. Also, the shift from coal use in the US might lead coal producers to look for markets overseas. Trade might have a significant impact on fuel prices, which we have already seen can greatly impact the US electricity sector, and vice versa (Chapter 5). The electricity-detailed representation in Chapter 5 was validated for the United States, but should be calibrated and validated for other regions. Integrating CGE linkages will unlock many timely research vistas including, but not limited to, the impact of different combinations of national and global emission policies as well as the impacts on fuel trade. This dissertation demonstrates the viability of electricity-detailed global economic analysis going forward.

REFERENCES

REFERENCES

- Aguiar, A., McDougall, R., and Narayanan, B. (2012). Global trade, assistance, and production: the GTAP 8 data base. *Center for Global Trade Analysis, Purdue University*.
- Apergis, N. and Payne, J. E. (2011). A dynamic panel study of economic development and the electricity consumption-growth nexus. *Energy Economics*, 33(5):770–781.
- Arndt, C. (1996). An introduction to systematic sensitivity analysis via gaussian quadrature. *GTAP Technical Papers*, 2.
- Arora, V. and Cai, Y. (2014). Disaggregating electricity generation technologies in CGE models.
- Bataille, C., Jaccard, M., Nyboer, J., and Rivers, N. (2006). Towards general equilibrium in a technology-rich model with empirically estimated behavioral parameters. *The Energy Journal*, pages 93–112.
- Beckman, J., Hertel, T., and Tyner, W. (2011). Validating energy-oriented CGE models. *Energy Economics*, 33(5):799–806.
- Bergman, L. (1991). General equilibrium effects of environmental policy: a CGE-modeling approach. *Environmental and Resource Economics*, 1(1):43–61.
- BLS, Bureau of Labor Statistics. (2015). Labor productivity and costs by industry tables. Technical report, Bureau of Labor Statistics.
- Böhringer, C. (1998). The synthesis of bottom-up and top-down in energy policy modeling. *Energy economics*, 20(3):233–248.

- Böhringer, C. and Rutherford, T. F. (2008). Combining bottom-up and top-down. *Energy Economics*, 30(2):574–596.
- Böhringer, C., Rutherford, T. F., and Tol, R. S. (2009). The EU 20/20/2020 targets: An overview of the EMF22 assessment. *Energy economics*, 31:S268–S273.
- Bowles, S. (1970). Aggregation of labor inputs in the economics of growth and planning: experiments with a two-level CES function. *The Journal of Political Economy*, pages 68–81.
- Brown, S. P. (1998). Global warming policy: some economic implications. *Economic Review-Federal Reserve Bank of Dallas*, pages 26–35.
- Burniaux, J.-M. and Truong, T. P. (2002). GTAP-E: an energy-environmental version of the GTAP model. *GTAP Technical Papers*, page 18.
- Bushnell, J. B. and Chen, Y. (2009). Regulation, allocation, and leakage in cap-and-trade markets for CO₂. Technical report, National Bureau of Economic Research.
- Bushnell, J. B., Holland, S. P., Hughes, J. E., and Knittel, C. R. (2014). Strategic policy choice in state-level regulation: The EPA’s Clean Power Plan. *University of California Energy Institute Working Paper*, 255.
- Butterfield, M. and Mules, T. (1980). A testing routing for evaluating cell by cell accuracy in short-cut regional input–output tables. *Journal of Regional Science*, 20(3):293–310.
- Cai, Y. and Arora, V. (2015). Disaggregating electricity generation technologies in CGE models: A revised technology bundle approach with an application to the US Clean Power Plan. *Applied Energy*, 154:543–555.
- Château, J., Dellink, R., and Lanzi, E. (2014). An overview of the OECD ENV-Linkages model.

- Chen, Y. (2009). Does a regional greenhouse gas policy make sense? a case study of carbon leakage and emissions spillover. *Energy Economics*, 31(5):667–675.
- Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *The American Economic Review*, 101(3):471–475.
- DeCanio, S. J. (2003). *Economic models of climate change: a critique*. Palgrave Macmillan New York.
- Delarue, E. and D'haeseleer, W. (2008). Greenhouse gas emission reduction by means of fuel switching in electricity generation: Addressing the potentials. *Energy Conversion and Management*, 49(4):843–853.
- Dixon, P. and Rimmer, M. (2002). *Dynamic general equilibrium modelling for forecasting and policy. A practical guide and documentation of MONASH*. Elsevier.
- Dixon, P. B. (1982). *ORANI, a multisectoral model of the Australian economy*, volume 142. North Holland.
- Dixon, P. B. and Rimmer, M. T. (2003). A new specification of labour supply in the MONASH model with an illustrative application. *Australian Economic Review*, 36(1):22–40.
- Dixon, P. B. and Rimmer, M. T. (2006). The displacement effect of labour-market programs: MONASH analysis*. *Economic Record*, 82(s1):S26–S40.
- Egging, R., Gabriel, S. A., Holz, F., and Zhuang, J. (2008). A complementarity model for the European natural gas market. *Energy policy*, 36(7):2385–2414.
- EIA (2015). International energy statistics.
- EIA, U.S. Energy Information Administration. (2014). Coal-fired power plant operators consider emissions compliance strategies. Technical report, U.S. Energy Information Administration - Independent Statistics and Analysis.

- EIA, U.S. Energy Information Administration. (2015a). Annual energy outlook 2015. Technical report, U.S. Energy Information Administration - Independent Statistics and Analysis.
- EIA, U.S. Energy Information Administration. (2015b). Electric power monthly. Technical report, U.S. Energy Information Administration - Independent Statistics and Analysis.
- EIA, U.S. Energy Information Administration. (2015c). NYMEX coal futures near-month contract final settlement price 2015. Technical report, U.S. Energy Information Administration.
- EPA, US Environmental Protection Agency. (2005). Response to comments for final Clean Air Mercury Rule - cost and energy impacts. Technical report, US Environmental Protection Agency.
- EPA, US Environmental Protection Agency. (2011). Regulatory impact analysis for the final Mercury and Air Toxics Standard. Technical report, US Environmental Protection Agency.
- EPA, US Environmental Protection Agency. (2015). Overview of the clean power plan: Cutting carbon pollution from power plants. Technical report, US Environmental Protection Agency.
- Ferguson, R., Wilkinson, W., and Hill, R. (2000). Electricity use and economic development. *Energy policy*, 28(13):923–934.
- Friedlander, D. (1961). A technique for estimating a contingency table, given the marginal totals and some supplementary data. *Journal of the Royal Statistical Society. Series A (General)*, pages 412–420.
- Giesecke, J. A., Tran, N. H., Corong, E. L., and Jaffee, S. (2013). Rice land designation policy in Vietnam and the implications of policy reform for food security and economic welfare. *The Journal of Development Studies*, 49(9):1202–1218.

- Golan, A., Judge, G., and Robinson, S. (1994). Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*, pages 541–549.
- Grubb, M., Edmonds, J., Ten Brink, P., and Morrison, M. (1993). The costs of limiting fossil-fuel CO₂ emissions: a survey and analysis. *Annual Review of Energy and the Environment*, 18(1):397–478.
- Haftendorn, C. and Holz, F. (2010). Modeling and analysis of the international steam coal trade. *The Energy Journal*, 31(4):205.
- Han, Z., Wei, Y., and Fan, Y. (2004). Research on change features of Chinese energy intensity and economic structure. *Application of Statistics and Management*, 1:001.
- Harrison, J., Horridge, M., Jerie, M., and Pearson, K. (2014). *GEMPACK manual*.
- Haszeldine, R. S. (2009). Carbon capture and storage: how green can black be? *Science*, 325(5948):1647–1652.
- Hazilla, M. and Kopp, R. J. (1990). Social cost of environmental quality regulations: A general equilibrium analysis. *Journal of Political Economy*, pages 853–873.
- Hirth, L., Ueckerdt, F., and Edenhofer, O. (2014). Why wind is not coal: on the economics of electricity.
- Hodge, B.-M. S., Huang, S., Sirola, J. D., Pekny, J. F., and Reklaitis, G. V. (2011). A multi-paradigm modeling framework for energy systems simulation and analysis. *Computers & Chemical Engineering*, 35(9):1725–1737.
- Horridge, M. (2005). SplitCom: Programs to disaggregate a GTAP sector. *Centre of Policy Studies, Monash University, Melbourne, Australia*.
- Hourcade, J.-C., Jaccard, M., Bataille, C., and Gherzi, F. (2006). Hybrid modeling: New answers to old challenges introduction to the special issue of “The Energy Journal”. *The Energy Journal*, pages 1–11.

- Huang, W., Kobayashi, S., and Tanji, H. (2008). Updating an input–output matrix with sign-preservation: some improved objective functions and their solutions. *Economic systems research*, 20(1):111–123.
- IEA, International Energy Agency. (2011). CO₂ emissions from fuel combustion-highlights. *IEA, Paris* <http://www.iea.org/co2highlights/co2highlights.pdf>. Cited July.
- IEA, International Energy Agency. (2014). Energy balances of non-OECD countries: Documentation for beyond 2020 files (2014 edition). Technical report, International Energy Agency.
- IEA, International Energy Agency. (2015). Energy balances of OECD countries: Documentation for beyond 2020 files (2015 edition). Technical report, International Energy Agency.
- IEA/NEA, International Energy Agency and OECD Nuclear Energy Agency. (2010). Projected costs of generating electricity - 2010 edition. Technical report, International Energy Agency and OECD Nuclear Energy Agency.
- Jaccard, M., Bailie, A., and Nyboer, J. (1996). CO₂ emission reduction costs in the residential sector: Behavioral parameters in a bottom-up simulation model. *The Energy Journal*, pages 107–134.
- Joskow, P. L. (1997). Restructuring, competition and regulatory reform in the US electricity sector. *The Journal of Economic Perspectives*, pages 119–138.
- Joskow, P. L. (2011). Comparing the costs of intermittent and dispatchable electricity generating technologies. *The American Economic Review*, pages 238–241.
- Junius, T. and Oosterhaven, J. (2003). The solution of updating or regionalizing a matrix with both positive and negative entries. *Economic Systems Research*, 15(1):87–96.

- Kiuiila, O. and Rutherford, T. (2013). Piecewise smooth approximation of bottom-up abatement cost curves. *Energy Economics*, 40:734–742.
- Knudsen, D. C. and Fotheringham, A. S. (1986). Matrix comparison, goodness-of-fit, and spatial interaction modeling. *International Regional Science Review*, 10(2):127–147.
- Koomey, J. G. (2005). Book review: Economic models of climate change: a critique. *Ecological Economics*, 55(1):131–132.
- Kozhevnikova, M. and Lange, I. (2009). Determinants of contract duration: further evidence from coal-fired power plants. *Review of Industrial Organization*, 34(3):217–229.
- Kullback, S. (1959). Information and statistics.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, pages 79–86.
- Kuroda, M. (1988). *A method of estimation for the updating transaction matrix in the input-output relationships*. Amsterdam: North Holland.
- Kuuskraa, V., Stevens, S. H., and Moodhe, K. D. (2013). Technically recoverable shale oil and shale gas resources: an assessment of 137 shale formations in 41 countries outside the United States. Technical report.
- Lahr, M. and De Mesnard, L. (2004). Biproportional techniques in input-output analysis: table updating and structural analysis. *Economic Systems Research*, 16(2):115–134.
- Lanz, B. and Rausch, S. (2011). General equilibrium, electricity generation technologies and the cost of carbon abatement: A structural sensitivity analysis. *Energy Economics*, 33(5):1035–1047.

- Lenzen, M. (2011). Aggregation versus disaggregation in input–output analysis of the environment. *Economic Systems Research*, 23(1):73–89.
- Lenzen, M., Wood, R., and Gallego, B. (2007). Some comments on the GRAS method. *Economic systems research*, 19(4):461–465.
- Levi, M. A. (2012). *A Strategy for US Natural Gas Exports*. Hamilton Project, Brookings Institution.
- Limmeechokchai, B. and Suksuntornsiri, P. (2007). Embedded energy and total greenhouse gas emissions in final consumptions within Thailand. *Renewable and Sustainable Energy Reviews*, 11(2):259–281.
- Lindner, S., Legault, J., and Guan, D. (2012). Disaggregating input–output models with incomplete information. *Economic Systems Research*, 24(4):329–347.
- Lindner, S., Legault, J., and Guan, D. (2013). Disaggregating the electricity sector of China’s input–output table for improved environmental life-cycle assessment. *Economic Systems Research*, 25(3):300–320.
- Loulou, R., Goldstein, G., Noble, K., et al. (2004). Documentation for the MARKAL family of models. *Energy Technology Systems Analysis Programme*, pages 65–73.
- Macmillan, S., Antonyuk, A., and Schwind, H. (2013). Gas to coal competition in the US power sector. Technical report, International Energy Agency, Insights Series 2013.
- Manne, A. S. and Richels, R. G. (1992). *Buying greenhouse insurance: the economic costs of carbon dioxide emission limits*. MIT press.
- Manne, A. S. and Wene, C.-O. (1992). MARKAL-MACRO: A linked model for energy-economy analysis. Technical report, Brookhaven National Lab., Upton, NY (United States).

- Marriott, J. (2007). *An electricity-focused economic input–output model: Life-cycle assessment and policy implications of future electricity generation scenarios*. PhD thesis, Carnegie Mellon University Pittsburgh, PA.
- McCollum, D., Bauer, N., Calvin, K., Kitous, A., and Riahi, K. (2014). Fossil resource and energy security dynamics in conventional and carbon-constrained worlds. *Climatic Change*, 123(3-4):413–426.
- McDougall, R. and Golub, A. (2008). A revised energy-environmental version of the GTAP model. *GTAP Technical Paper Series*.
- McDougall, R. A. (1992). Flexibly nested production functions: implementation for MONASH. Technical report, Victoria University, Centre of Policy Studies/IMPACT Centre.
- McDougall, R. A. (1999). Entropy theory and ras are friends. *GTAP Working Papers*, 6.
- McFarland, J. R., Reilly, J. M., and Herzog, H. J. (2004). Representing energy technologies in top-down economic models using bottom-up information. *Energy Economics*, 26(4):685–707.
- Mínguez, R., Oosterhaven, J., and Escobedo, F. (2009). Cell-corrected RAS method (CRAS) for updating or regionalizing an input–output matrix. *Journal of Regional Science*, 49(2):329–348.
- Monitoring Analytics (2015). State of the market report for PJM.
- Neij, L. (2008). Cost development of future technologies for power generation - a study based on experience curves and complementary bottom-up assessments. *Energy policy*, 36(6):2200–2211.
- Obama, B. (2013). Presidential memorandum – power sector carbon pollution standards. memorandum to the environmental protection agency. *The White House Office of the Press Secretary*.

- Paltsev, S., Reilly, J. M., Jacoby, H. D., Eckaus, R. S., McFarland, J. R., Sarofim, M. C., Asadoorian, M. O., and Babiker, M. H. (2005). The MIT emissions prediction and policy analysis (EPPA) model: version 4. Technical report, MIT Joint Program on the Science and Policy of Global Change.
- Pant, H. (2007). GTEM: global trade and environment model. *Australian Bureau of Agricultural and Resource Economics, Canberra*.
- Payne, J. E. (2010). A survey of the electricity consumption-growth literature. *Applied Energy*, 87(3):723–731.
- Peters, G. P. and Hertwich, E. G. (2008). CO₂ embodied in international trade with implications for global climate policy. *Environmental Science & Technology*, 42(5):1401–1407.
- Peters, J. C. (2015). The GTAP-Power database: Disaggregating the electricity sector in the GTAP database. *GTAP Resource*, 4619.
- Peters, J. C. and Hertel, T. W. (2015a). The database-modeling nexus in integrated assessment modeling of electric power generation. *In Review*.
- Peters, J. C. and Hertel, T. W. (2015b). Matrix balancing with unknown total costs: An example in the electric power disaggregation. *In Review*.
- Quadrelli, R. and Peterson, S. (2007). The energy–climate challenge: recent trends in CO₂ emissions from fuel combustion. *Energy Policy*, 35(11):5938–5952.
- Robinson, S., Cattaneo, A., and El-Said, M. (2001). Updating and estimating a social accounting matrix using cross entropy methods. *Economic Systems Research*, 13(1):47–64.
- Rutherford, T. F. (2005). Book review: Economic models of climate change: A critique. *Journal of Economic Literature*, 43(2):525–525.

- Schäfer, A. and Jacoby, H. D. (2005). Technology detail in a multisector CGE model: transport under climate policy. *Energy Economics*, 27(1):1–24.
- Schäfer, A. and Jacoby, H. D. (2006). Experiments with a hybrid CGE-MARKAL model. *The Energy Journal*, pages 171–177.
- Schneider, M. H. and Zenios, S. A. (1990). A comparative study of algorithms for matrix balancing. *Operations research*, 38(3):439–455.
- Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423,623–656.
- Shrestha, R. M. and Marpaung, C. O. (2006). Integrated resource planning in the power sector and economy-wide changes in environmental emissions. *Energy Policy*, 34(18):3801–3811.
- Smith, D. and Hutchinson, B. (1981). Goodness of fit statistics for trip distribution models. *Transportation Research Part A: General*, 15(4):295–303.
- Soytas, U., Sari, R., and Ewing, B. T. (2007). Energy consumption, income, and carbon emissions in the United States. *Ecological Economics*, 62(3):482–489.
- Stuermer, M. and Schwerhoff, G. (2013). Technological change in resource extraction and endogenous growth. *Bonn Econ Discussion Papers*, 12.
- Sue Wing, I. (2008). The synthesis of bottom-up and top-down approaches to climate policy modeling: Electric power technology detail in a social accounting framework. *Energy Economics*, 30(2):547–573.
- Temurshoev, U. (2012). *Entropy-based benchmarking methods*. University of Groningen, Groningen Growth and Development Centre.
- Temurshoev, U., Webb, C., and Yamano, N. (2011). Projection of supply and use tables: Methods and their empirical assessment. *Economic systems research*, 23(1):91–123.

- Tidball, R., Bluestein, J., Rodriguez, N., and Knoke, S. (2010). Cost and performance assumptions for modeling electricity generation technologies. Technical Report NREL/SR-6A20-48595, National Renewable Energy Laboratory.
- United Nations (2012). UN commodity trade statistics database.
- van der Mensbrugge, D. and Peters, J. C. (2015). Quantity-preserving constant elasticity of substitution in electric power generation. *In Preparation*.
- Wang, Y. and Brown, M. A. (2014). Policy drivers for improving electricity end-use efficiency in the usa: an economic–engineering analysis. *Energy Efficiency*, 7(3):517–546.
- Weisser, D. (2007). A guide to life-cycle greenhouse gas (GHG) emissions from electric supply technologies. *Energy*, 32(9):1543–1559.
- West, G. R. (1995). Comparison of input–output, input–output + econometric and computable general equilibrium impact models at the regional level. *Economic Systems Research*, 7(2):209–227.
- Williams, J. H., DeBenedictis, A., Ghanadan, R., Mahone, A., Moore, J., Morrow, W. R., Price, S., and Torn, M. S. (2012). The technology path to deep greenhouse gas emissions cuts by 2050: the pivotal role of electricity. *Science*, 335(6064):53–59.
- Wing, I. S. (2006). The synthesis of bottom-up and top-down approaches to climate policy modeling: electric power technologies and the cost of limiting US CO₂ emissions. *Energy Policy*, 34(18):3847–3869.
- Wiskich, A. (2014). Implementing a load duration curve of electricity demand in a general equilibrium model. *Energy Economics*, 45:373–380.
- World Bank (2012). *World development indicators 2012*. World Bank Publications.
- Yunfeng, Y. and Laike, Y. (2010). China’s foreign trade and climate change: a case study of CO₂ emissions. *Energy Policy*, 38(1):350–356.

APPENDICES

Appendix A. Capacity Flexibility, Utilization, and Expansion Analytics: Technology Price Shock

Let us first consider a simple model which characterizes capacity flexibility, utilization, expansion, and their interactions. First, we have percentage change in total electricity demand, \widehat{Q}^g , determined by elasticity of demand, η_D , and the percentage change in aggregate cost of electricity, \widehat{p}^g :

$$\widehat{Q}^g = -\eta_D \cdot \widehat{p}^g \quad (\text{A.1})$$

The aggregate cost of electricity is subject to the following zero-profit or aggregate cost pricing condition:

$$\widehat{p}^g = \sum_t \theta_t^g \widehat{p}_t^g \quad (\text{A.2})$$

where θ_t^g is the share of total electricity production from technology t in GWh, and \widehat{p}_t^g is the percentage change in total cost of electricity generated using technology t . Total cost of electricity using technology t is also given by a zero-profit or aggregate cost pricing condition:

$$\widehat{p}_t^g = \sum_i \theta_{it} \widehat{p}_{it} + \delta_t^p \quad (\text{A.3})$$

where θ_{it} is the cost share of input i in producing electricity with technology t and \widehat{p}_{it} is the percentage change in input cost. Here, the cost of generating electricity using technology t is subject to a price shock, δ_t^p . This will allow us to determine the effect of parameter values on the responsiveness of the returns to capacity given an exogenous price shock.

The derived demand equation allows for substitution between input prices given by the constant elasticity of substitution between inputs, σ_i . A flexible technology has $\sigma_i > 0$ and an inflexible technology has $\sigma_i = 0$.

$$\widehat{q}_{it}^D = \widehat{q}_t^g - \sigma_i (\widehat{p}_{it} - \widehat{p}_t^g) \quad (\text{A.4})$$

where \widehat{q}_t^g is the percentage change in total generation using technology t . This is given by the corresponding derived demand equation that allows for substitution between technologies using existing capacity.

$$\widehat{q}_t^g = \widehat{Q}^g - \sigma_t (\widehat{p}_t^g - \widehat{p}^g) \quad (\text{A.5})$$

where σ_t is a CES parameter that characterizes the ability of the electricity sector to adjust generation via the utilization mechanism.

Finally, we represent the capacity expansion mechanism with the following elasticity of supply, μ_t :

$$\widehat{q}_{kt}^S = \mu_t \cdot \widehat{p}_{kt} \quad (\text{A.6})$$

To make the system of equations analytically tractable, we assume that only change in prices is in capital inputs of the technology that we are investigating (i.e infinite supply elasticity for other inputs and technologies). That is,

$$\widehat{p}_{it} = 0 \text{ for all } i \neq k \quad (\text{A.7})$$

and

$$\widehat{p}_t^g = 0 \text{ for all } t \neq s \quad (\text{A.8})$$

where capital is designated k and the technology we are investigating is designated s . Therefore, we rewrite the original zero-profit/average cost pricing equations (Equations A.3 and A.2) as:

$$\widehat{p}_s^g = \theta_{ks} \widehat{p}_{ks} + \delta_s^p \quad (\text{A.9})$$

$$\widehat{p}^g = \theta_s^g \theta_{ks} \widehat{p}_{ks} + \theta_s^g \delta_s^p \quad (\text{A.10})$$

This is a simplification of the electricity-detailed computational equilibrium model; however, using this analysis, we can show the interdependencies between capacity utilization and expansion.

We begin by equating capacity supply and demand for technology s using Equations A.6 and A.4).

$$\mu_s \widehat{p}_{ks} = \widehat{q}_{ks}^S = \widehat{q}_{ks}^D = \widehat{q}_s^g - \sigma_i (\widehat{p}_{ks} - \widehat{p}_s^g) \quad (\text{A.11})$$

Simplifying and distributing σ_i nets

$$\mu_s \widehat{p}_{ks} = \widehat{q}_s^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{A.12})$$

and substituting \widehat{q}_s^g using Equation A.5 allows us to write:

$$\mu_s \widehat{p}_{ks} = \widehat{Q}^g - \sigma_t \widehat{p}_s^g + \sigma_t \widehat{p}^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{A.13})$$

Replacing \widehat{Q}^g using Equation A.1 nets the following equation with only price percentage change variables and parameters.

$$\mu_s \widehat{p}_{ks} = -\eta_D \widehat{p}^g - \sigma_t \widehat{p}_s^g + \sigma_t \widehat{p}^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{A.14})$$

Substituting with Equations A.9 and A.10 allows us to write the relationship in terms of only the parameters, price shock, and returns to capacity:

$$\begin{aligned} \mu_s \widehat{p}_{ks} = & -\eta_D \theta_s^g \theta_{ks} \widehat{p}_{ks} - \eta_D \theta_s^g \delta_s^p - \sigma_t \theta_{ks} \widehat{p}_{ks} - \sigma_t \delta_s^p \\ & + \sigma_t \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_t \theta_s^g \delta_s^p - \sigma_i \widehat{p}_{ks} + \sigma_i \theta_{ks} \widehat{p}_{ks} + \sigma_i \delta_s^p \end{aligned} \quad (\text{A.15})$$

Collecting all the \widehat{p}_{ks} on the left hand side and $\widehat{\delta}_s^p$ on the right hand side leads to the following equations:

$$\begin{aligned}
& \mu_s \widehat{p}_{ks} + \eta_D \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_t \theta_{ks} \widehat{p}_{ks} - \sigma_t \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_i \widehat{p}_{ks} - \sigma_i \theta_{ks} \widehat{p}_{ks} \\
& = -\eta_D \theta_s^g \delta_s^p - \sigma_t \delta_s^p + \sigma_t \theta_s^g \delta_s^p + \sigma_i \delta_s^p
\end{aligned} \tag{A.16}$$

$$\widehat{p}_{ks} (\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t \theta_{ks} - \sigma_t \theta_s^g \theta_{ks} + \sigma_i - \sigma_i \theta_{ks}) = -\delta_s^p (\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i) \tag{A.17}$$

This leads us to define the response in returns to capacity, \widehat{p}_{ks} , as a function of the price shock, $-\delta_s^p$, and relationships between the parameters. The following three equations are equivalent; however, the parameters are rearranged to highlight various relationships in the parameters which affect the response of the returns to capacity in response to the price shock.

$$\widehat{p}_{ks} = -\delta_s^p \cdot \frac{\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i}{\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t \theta_{ks} - \sigma_t \theta_s^g \theta_{ks} + \sigma_i - \sigma_i \theta_{ks}} \tag{A.18}$$

$$\widehat{p}_{ks} = -\delta_s^p \cdot \frac{\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i}{\mu_s + \theta_{ks} \cdot (\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i) + \sigma_i} \tag{A.19}$$

$$\widehat{p}_{ks} = -\delta_s^p \cdot \frac{\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i}{\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t (\theta_{ks} - \theta_s^g \theta_{ks}) + \sigma_i \cdot (1 - \theta_{ks})} \tag{A.20}$$

It is easy to see that for $\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i > 0$ a negative price shock (i.e. cheaper generation with technology s) results in higher returns to capacity. We are able to draw out the following insights from the parameter terms in Equations A.18–A.20. Recall that σ_i relates to the flexibility of the technology, σ_t to capacity utilization, and μ_s to capacity expansion.¹

1. If μ_s increases (i.e. higher capacity expansion elasticity via MNL) then the response of \widehat{p}_{ks} to δ_s^p decreases (i.e. denominator in Equation A.18 increases). In other words, returns to capacity is dampened by responsive capacity expansion (e.g. low rate of retirement for the technology, high rate of total retirements,

¹The following assumes that $\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i > 0$.

high MNL parameter). This is one of the major linkages between utilization and expansion.

2. If σ_i (i.e. flexibility of the technology) increases then the response of \widehat{p}_{ks} to δ_s^p decreases. Equation A.20 shows the numerator decreases and the denominator increases. The response of cost of capacity is offset by substituting away from capital by increasing labor and operations.
3. If there is greater opportunity to increase capacity utilization (i.e. σ_t increases), returns to capacity will increase because there will be more generation per unit of capacity. The numerator increases more than the denominator (Equation A.20).
4. If η_D increases then the response of \widehat{p}_{ks} to δ_s^p increases. The numerator increases more than the denominator (Equation A.20).
5. The greater the capital intensity of the technology (i.e. θ_{ks}), the less the response of \widehat{p}_{ks} to δ_s^p . The denominator increases more than the numerator in Equation A.19.
6. The impact of the share of the technology in the generation mix (i.e. θ_s^g) depends on the relationship of $(\eta_D - \sigma_t)$ (see corresponding terms in Equation A.19). With high opportunity for capacity utilization changes ($\sigma_t > \eta_D$), the greater the share of the technology in the mix, the greater the response in returns to capacity.

Appendix B. Capacity Flexibility, Utilization, and Expansion Analytics: Electricity Demand Shock

Let us first consider a simple model which characterizes capacity flexibility, utilization, expansion, and their interactions. First, we have percentage change in total electricity demand, \widehat{Q}^g , determined by elasticity of demand, η_D , and the percentage change in aggregate cost of electricity, \widehat{p}^g :

$$\widehat{Q}^g = -\eta_D \cdot \widehat{p}^g + \delta^D \quad (\text{B.1})$$

The aggregate cost of electricity is subject to the following zero-profit or aggregate cost pricing condition:

$$\widehat{p}^g = \sum_t \theta_t^g \widehat{p}_t^g \quad (\text{B.2})$$

where θ_t^g is the share of total electricity production from technology t in GWh, and \widehat{p}_t^g is the percentage change in total cost of electricity generated using technology t . Total cost of electricity using technology t is also given by a zero-profit or aggregate cost pricing condition:

$$\widehat{p}_t^g = \sum_i \theta_{it} \widehat{p}_{it} \quad (\text{B.3})$$

where θ_{it} is the cost share of input i in producing electricity with technology t and \widehat{p}_{it} is the percentage change in input cost. Here, the cost of generating electricity using technology t is subject to a demand shock, δ_t^p . This will allow us to determine the effect of parameter values on the responsiveness of the returns to capacity given an exogenous demand shock.

The derived demand equation allows for substitution between input prices given by the constant elasticity of substitution between inputs, σ_i . A flexible technology has $\sigma_i > 0$ and an inflexible technology has $\sigma_i = 0$.

$$\widehat{q}_{it}^D = \widehat{q}_t^g - \sigma_i (\widehat{p}_{it} - \widehat{p}_t^g) \quad (\text{B.4})$$

where \widehat{q}_t^g is the percentage change in total generation using technology t . This is given by the corresponding derived demand equation that allows for substitution between technologies using existing capacity.

$$\widehat{q}_t^g = \widehat{Q}^g - \sigma_t (\widehat{p}_t^g - \widehat{p}^g) \quad (\text{B.5})$$

where σ_t is a CES parameter that characterizes the ability of the electricity sector to adjust generation via the utilization mechanism.

Finally, we represent the capacity expansion mechanism with the following elasticity of supply, μ_t :

$$\widehat{q}_{kt}^S = \mu_t \cdot \widehat{p}_{kt} \quad (\text{B.6})$$

To make the system of equations analytically tractable, we assume that only change in prices is in capital inputs of the technology that we are investigating (i.e infinite supply elasticity for other inputs and technologies). That is,

$$\widehat{p}_{it} = 0 \text{ for all } i \neq k \quad (\text{B.7})$$

where capital is designated k and the technology we are investigating is designated s . Therefore, we rewrite the original zero-profit/average cost pricing equations (Equations B.3 and B.2) as:

$$\widehat{p}_s^g = \theta_{ks} \widehat{p}_{ks} \quad (\text{B.8})$$

$$\widehat{p}^g = \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt} \quad (\text{B.9})$$

This is a simplification of the electricity-detailed computational equilibrium model; however, using this analysis, we can show the interdependencies between capacity utilization and expansion.

We begin by equating capacity supply and demand for technology s using Equations B.6 and B.4).

$$\mu_s \widehat{p}_{ks} = \widehat{q}_{ks}^S = \widehat{q}_{ks}^D = \widehat{q}_s^g - \sigma_i (\widehat{p}_{ks} - \widehat{p}_s^g) \quad (\text{B.10})$$

Simplifying and distributing $\widehat{\sigma}_i$ nets

$$\mu_s \widehat{p}_{ks} = \widehat{q}_s^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{B.11})$$

and substituting \widehat{q}_s^g using Equation B.5 allows us to write:

$$\mu_s \widehat{p}_{ks} = \widehat{Q}^g - \sigma_t \widehat{p}_s^g + \sigma_t \widehat{p}^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{B.12})$$

Replacing \widehat{Q}^g using Equation B.1 nets the following equation with only price percentage change variables and parameters.

$$\mu_s \widehat{p}_{ks} = \delta^D - \eta_D \widehat{p}^g - \sigma_t \widehat{p}_s^g + \sigma_t \widehat{p}^g - \sigma_i \widehat{p}_{ks} + \sigma_i \widehat{p}_s^g \quad (\text{B.13})$$

Substituting with Equations B.8 and B.9 allows us to write the relationship in terms of only the parameters, demand shock, and returns to capacity:

$$\begin{aligned} \mu_s \widehat{p}_{ks} = & \delta^D - \eta_D \theta_s^g \theta_{ks} \widehat{p}_{ks} - \eta_D \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt} - \sigma_t \theta_{ks} \widehat{p}_{ks} \\ & + \sigma_t \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_t \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt} - \sigma_i \widehat{p}_{ks} + \sigma_i \theta_{ks} \widehat{p}_{ks} \end{aligned} \quad (\text{B.14})$$

Collecting all the \widehat{p}_{ks} on the left hand side and $\widehat{\delta}_s^p$ on the right hand side leads to the following equations:

$$\begin{aligned}
& \mu_s \widehat{p}_{ks} + \eta_D \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_t \theta_{ks} \widehat{p}_{ks} - \sigma_t \theta_s^g \theta_{ks} \widehat{p}_{ks} + \sigma_i \widehat{p}_{ks} - \sigma_i \theta_{ks} \widehat{p}_{ks} \\
& = \delta^D - \eta_D \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt} + \sigma_t \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt}
\end{aligned} \tag{B.15}$$

$$\widehat{p}_{ks} (\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t \theta_{ks} - \sigma_t \theta_s^g \theta_{ks} + \sigma_i - \sigma_i \theta_{ks}) = \delta^D + (\sigma_t - \eta_D) \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt} \tag{B.16}$$

This leads us to define the response in returns to capacity, \widehat{p}_{ks} , as a function of the demand shock, δ^D , and relationships between the parameters. The following three equations are equivalent; however, the parameters are rearranged to highlight various relationships in the parameters which affect the response of the returns to capacity in response to the demand shock. Not that the denominator is identical to the denominator in Appendix A which explored a technology-specific price shock.

$$\widehat{p}_{ks} = \frac{\delta^D + (\sigma_t - \eta_D) \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt}}{\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t \theta_{ks} - \sigma_t \theta_s^g \theta_{ks} + \sigma_i - \sigma_i \theta_{ks}} \tag{B.17}$$

$$\widehat{p}_{ks} = \frac{\delta^D + (\sigma_t - \eta_D) \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt}}{\mu_s + \theta_{ks} \cdot (\eta_D \theta_s^g + \sigma_t - \sigma_t \theta_s^g - \sigma_i) + \sigma_i} \tag{B.18}$$

$$\widehat{p}_{ks} = \frac{\delta^D + (\sigma_t - \eta_D) \sum_{t \neq s} \theta_t^g \theta_{kt} \widehat{p}_{kt}}{\mu_s + \eta_D \theta_s^g \theta_{ks} + \sigma_t (\theta_{ks} - \theta_s^g \theta_{ks}) + \sigma_i \cdot (1 - \theta_{ks})} \tag{B.19}$$

VITA

VITA

Autobiographical information on Jeffrey C. Peters including up-to-date projects, publications, and personal details is available at: <http://www.jeffreypeters.com>