A METHOD FOR STAKEHOLDER-BASED COMPARATIVE BENCHMARKING OF AIRPORTS

by

Claes Johan David Schaar
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Information Technology

Committee:

	Dr. Lance Sherry, Dissertation Director
	Dr. Alexander Brodsky, Committee Member
	Dr. George Donohue, Committee Member
	Dr. John Shortle, Committee Member
	Dr. Daniel Menascé, Senior Associate Dean
	Dr. Lloyd J. Griffiths, Dean, The Volgenau School of Information Technology and Engineering
Date: 7/30 2010	Summer Semester 2010 George Mason University Fairfax, VA

A Method for Stakeholder-Based Comparative Benchmarking of Airports

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Claes Johan David Schaar
Master of Science
Linköping Institute of Technology, 2003

Director: Dr. Lance Sherry, Professor
Department of Systems Engineering and Operations Research

Summer Semester 2010
George Mason University
Fairfax, VA

DEDICATION

To my father and grandfather for the inspiration to begin this work and to Aimee for her encouragement, patience, and belief in me, without which I couldn't have completed it.

ACKNOWLEDGEMENTS

I would like to acknowledge the role of Dr. Lance Sherry who was my advisor in completing this dissertation. His ability to help me elevate my level of thinking about the problem, his knowledge of the domain, and his persistent encouragement throughout every part of the process were essential to my completion of the dissertation. I would like to express my deep gratitude to him. I also want to thank the members of my dissertation committee, Dr. Alexander Brodsky, Dr. George Donohue, and Dr. John Shortle, for their insights in helping set the direction of my work.

My sincere thanks go to the 32 people who generously shared their time with me and provided expert knowledge about airport stakeholders, to the individuals who introduced me to these subject-matter experts, and to those who reviewed my work and provided insightful feedback and directions.

I want to express my appreciation to my fellow students at the Center for Air Transportation System Research for their help with many technical and data-related questions and for their friendship during the research process. Thanks also to Lisa Nolder of the Graduate Student Services office, whose help was invaluable in navigating the many steps of life as a graduate student.

Finally, and most of all, I want to thank my wife Aimee. Her patience and support have carried me through this process. She had to make many sacrifices to allow me to pursue this work, and for that I am always grateful.

TABLE OF CONTENTS

LIST OF TABLES.....xi

LIST OF FIGURES.....xv

LIST OF ABBREVIATIONS.....xxii

ABSTRACT.....xxv

1 CHAPTER 1: INTRODUCTION..... 6

1.1 Scope of the Dissertation..... 9

1.2 Problem Statement and Summary of Unique Contributions..... 13

1.2.1 Problem Statement..... 13

1.2.2 Unique Contributions of Dissertation..... 14

1.3 Summary of Dissertation Method..... 16

1.4 Summary of Results..... 19

2 CHAPTER 2: A REVIEW OF AIRPORT BENCHMARKING..... 24

2.1 The Airport as a Public Utility..... 24

2.1.1 Public Utilities..... 24

2.1.2 Airport Finance..... 26

2.1.3 A Model of U.S. Airport Stakeholders and their Goals for the Airport ... 30

2.2 Airport Benchmarking..... 69

2.2.1 Overview of Benchmarking..... 70

2.2.2 Analytical Techniques for Benchmarking..... 73

2.2.3 Review of Airport Benchmarks 103

2.2.4 Implications of the Stakeholder Analysis on Airport Performance Benchmarking 125

2.2.5 Analytical Techniques Used in Past Airport Benchmarks 127

2.3 Dissertation Problem Statement 130

2.3.1 Problem 1: Stakeholder and Goal Ambiguity 130

2.3.2 Problem 2: Lack of Systematic Model Selection 131

2.3.3 Problem 3: No Benchmarks Apply a Systematic Process 131

3 CHAPTER 3: METHODOLOGY 133

3.1 Step 1: Select Stakeholders and Goals 137

3.2 Step 2: Identify Metrics for Goals 139

3.2.1 Selecting the Metrics 139

3.2.2 Identifying Methodological Discrepancies 143

3.3 Step 3: Select Benchmark Model 144

3.3.1 A DEA Model Selection Framework 145

3.3.2 Heuristics for Making Choices in the DEA Model Selection Framework 147

3.4 Step 4: Collect Data 153

3.5 Step 5: Compute Benchmark Results 155

3.5.1 Matlab Implementation 157

3.5.2 C++ Implementation 157

3.6 Step 6: Interpret Results and Derive Findings 158

3.6.1 Identifying Factors which Impact Benchmark Results 159

3.6.2 Categorizing Airports According to Environmental Factors 161

3.6.3 Investigating Individual Airports 162

4 CHAPTER 4: RESULTS 163

4.1 Assessment of Validity of Past Benchmarks: Application of the DEA Model Selection Framework and Heuristics to Past Airport Benchmarks 164

4.1.1 Results of Framework and Heuristics Application 164

4.1.2 Areas of Discrepancy between Study Choices and Heuristics..................181

4.1.3 Implications of the Gaps between Study Choices and Heuristics 185

4.2 Case Study 1: Benchmarking the Level of Domestic Air Service to U.S. Metropolitan Areas 187

4.2.1 The Airport's Stakeholders and Their Goals 188

4.2.2 Methodology 190

4.2.3 Results..................218

4.2.4 Conclusions 251

4.3 Case Study 2: Benchmark of the Level of Capacity Utilization at U.S. Airports 252

4.3.1 The Airport's Stakeholders and Their Goals 255

4.3.2 Methodology 257

4.3.3 Results..................281

4.3.4 Conclusions 314

4.4 Case Study 3: Re-design of an Existing Benchmark 315

4.4.1 Original Study..................318

4.4.2 Design of New Study..................334

4.4.3 Data, Results, and Discussion of New Study..................359

4.4.4 Comparison of Original and New Benchmark Results 391
4.4.5 Conclusions 396

5 CHAPTER 5: CONCLUSIONS AND FUTURE WORK 398
5.1 Airports as Utilities 398
5.2 Limitations of Benchmarking Methods 400
5.3 A Comprehensive Benchmarking Methodology 402
5.4 Implications for Airport Policymakers 404
5.4.1 Summary of Benchmark Results 404
5.4.2 Use of Benchmarks by Policymakers 407
5.4.3 Implications for Decisions about Funding 407
5.4.4 Motivations of Airport Management 417
5.4.5 Airports' Roles in their Regions 420
5.4.6 Conflicting Objectives among Airport Stakeholders 422
5.5 Future Work 423

APPENDIX A: A FRAMEWORK AND HEURISTICS FOR DEA MODEL SELECTION IN AIRPORT BENCHMARKING 427
A.1 Existing Frameworks for Analysis of DEA Models 428
A.2 Extending a Framework for Selection of DEA Models 433
A.2.1 Aggregation 437
A.2.2 Weights 441
A.2.3 Orientation 442
A.2.4 Returns to Scale 444
A.2.5 Free Disposal Hull 445

A.2.6 Integer Constraints 446

A.2.7 Time Horizon 447

A.2.8 Tie-Breaking 447

A.3 Developing Heuristics for DEA Modeling of Airport Performance 449

A.3.1 Aggregation 450

A.3.2 Weights 452

A.3.3 Orientation 453

A.3.4 Returns to Scale 453

A.3.5 Free Disposal Hull 455

A.3.6 Integer Constraints 456

A.3.7 Time Horizons 456

A.3.8 Tie-Breaking 458

A.3.9 Summary of Airport DEA Model Selection Heuristics 459

APPENDIX B: MATLAB DEA CODE 465

B.1 Code for Single DEA Run 465

B.2 Code for Batch Execution of DEA Run 482

APPENDIX C: C++ DEA CODE 483

APPENDIX D: STAKEHOLDER KNOWLEDGE ELICITATION 497

D.1 Subject-Matter Experts 497

D.2 Knowledge Elicitation Form 501

APPENDIX E: DETAILED BENCHMARK DATA AND RESULTS 502

E.1 Case Study 1: Benchmarking the Level of Domestic Air Service to U.S. Metropolitan Areas 504

E.1.1 Benchmark Data 504

E.1.2 Results 508

E.2 Case Study 2: Benchmark of the Level of Capacity Utilization at U.S. Airports 509

E.2.1 Benchmark Data 509

E.2.2 Results 513

E.3 Case Study 3: Re-design of an Existing Benchmark 514

E.3.1 Benchmark Data 514

E.3.2 Results 523

REFERENCES 526

CURRICULUM VITAE 544

LIST OF TABLES

Table 1.1 - OEP-35 airports (FAA 2009) ... 9

Table 1.2 - Unique contributions of dissertation ... 15

Table 2.1 - Airport stakeholders ... 31

Table 2.2 – Description of airport stakeholders and goals ... 57

Table 2.3 - System loops in stakeholder model ... 64

Table 2.4 - Determination of SXEF (Doyle & Green 1994) ... 95

Table 2.5 - Overview of past airport benchmarks ... 105

Table 2.6 - Industry Airport Benchmarks ... 121

Table 2.7 - Summary of metrics used in industry benchmarks ... 124

Table 2.8 - Models used in past airport benchmark studies ... 128

Table 3.1 - Phases of benchmarking methodology ... 134

Table 3.2 - Development of benchmarking methodology ... 135

Table 3.3 – Heuristics for airport benchmarking using the DEA model choice framework. Full details available in Appendix A. ... 147

Table 3.4 - Translation of heuristics to specific model choices ... 150

Table 3.5 - Overview of DEA software ... 155

Table 4.1 - Results of application of DEA framework and heuristics to past academic airport benchmarks ... 166

Table 4.2 - Application of DEA framework and heuristics to past airport benchmark studies ... 169

Table 4.3 - Airports included in study 191

Table 4.4 - DEA model parameter choices 203

Table 4.5 - Non-OEP-35 airports added to the study 206

Table 4.6 - Average number of domestic airports served nonstop at least 52 times annually (source: T100 database) 210

Table 4.7 - Average number of OEP-35 airports served nonstop at least 52 times annually (source: T100 database) 210

Table 4.8 – Distance to the air service frontier. These are gaps in the level of service to be closed for achieving air service level of 1.00. The gaps are the shortest distance to the frontier. 222

Table 4.9 - Summary of study results, 2005-2008, in order of GDP per capita. Areas with level of air service scores above 1.3 are highlighted in yellow as those areas have poor levels of air service. 225

Table 4.10 - Planned airport capacity improvements (Federal Aviation Administration 2007). For greyed-out airports no data was available. Capacity improvements that apply to all OEP-35 airports have been omitted 232

Table 4.11 - Definition of NAICS categories included in high-tech industry (U.S. Census Bureau 2002) 236

Table 4.12 - Results of Kruskal-Wallis test of the benchmark results based on grouping 240

Table 4.13 - Rankings from selected sensitivity tests. Test 1 has the least restrictive weight boundaries and Test 7 has the most restrictive boundaries. 247

Table 4.14 – Sample measures of airport capacity used in past benchmarks 262

Table 4.15 - DEA model parameter choices 268

Table 4.16 - Average number of domestic airports served nonstop at least 52 times annually (source: T100 database) 272

Table 4.17 - Average number of OEP-35 airports served nonstop at least 52 times annually (source: T100 database) 272

Table 4.18 - Portion of all flights that were domestic passenger flights 276

Table 4.19 - Distance to the capacity utilization frontier. These are gaps in the outputs to be closed for achieving a score of 1.00. The gaps are the distance to the frontier. 299

Table 4.20 - Rankings from selected sensitivity tests. Test 1 has the least restrictive weight boundaries and Test 7 has the most restrictive boundaries 303

Table 4.21 - Data sources for original case study 322

Table 4.22 - DEA model choices in original study 330

Table 4.23- Kruskal-Wallis test on benchmark result rankings 333

Table 4.24 - Categorization of performance metrics in original study 335

Table 4.25 - Mapping of goals to stakeholder for new case study 337

Table 4.26 - Pearson correlation coefficients for test of correlation between delay metrics. The delay data was computed annually for 2005-2008 for the OEP-35 airports, and the total delay for all flights was summed for each year 342

Table 4.27 - Factors considered in airport credit ratings (Forsgren 2007). Categorizations added by this author. 344

Table 4.28 - Regional growth factors 346

Table 4.29 - Air service factors 346

Table 4.30 - Airfield/facilities factors 348

Table 4.31 - Financial factors 349

Table 4.32 - Modeling choices for benchmark 356

Table 4.33 - Modeling choices for benchmark 359

Table 4.34 – Results from Kruskal-Wallis test of operating efficiency benchmark results ranking based on airport size 367

Table 4.35 - Results from Kruskal-Wallis test of operating efficiency benchmark results ranking based on weather conditions 370

Table 4.36 - Results from Kruskal-Wallis test of air service benchmark results ranking for small and large airports 383

Table 4.37 - Results from Kruskal-Wallis test of financial factors benchmark results ranking for small and large airports 383

Table 4.38 - Results from Kruskal-Wallis test of financial factors benchmark results ranking for low and high degrees of connecting traffic. 386

Table 4.39 - Credit rating of most senior airport liens (Lehman et al. 2009)....387

Table 4.40 - Spearman rank correlation coefficients for test of correlation between benchmark result rankings....393

Table A.1 - A DEA classification framework (Kleine 2004) 429

Table A.2 - Structure of a DEA model framework for airport benchmarking. Each section in the framework describes an element for which the DEA modeler must make a selection from the choices presented....435

Table A.3 - Airport DEA Framework and Heuristics....459

Table A.4 - Translation of heuristics to specific model choices....462

Table D.1 – Participants in knowledge elicitation sessions on the definition of airport stakeholders, their goals, and key performance indicators for airports....497

Table E.1 - Case study 1 benchmark data....504

Table E.2 - Case study 1 benchmark results....508

Table E.3 - Case study 2 benchmark data....509

Table E.4 - Case study 2 benchmark results....513

Table E.5 - Case study 3 benchmark data for original study....514

Table E.6 - Case study 3 benchmark data for operational efficiency study 518

Table E.7 - Case study 3 benchmark data for investment quality study 522

Table E.8 - Case study 3 benchmark results for original study 523

Table E.9 - Case study 3 benchmark results for operational efficiency study 524

Table E.10 - Case study 3 benchmark results for investment quality study 525

LIST OF FIGURES

Figure 1.1 – Positively reinforcing relationship between airport growth and regional economic growth 7

Figure 1.2 - Overview of the benchmarking methodology 17

Figure 2.1 - Airport capital funding sources for large hubs (Airports Council International - North America 2009a, p. 10) 28

Figure 2.2 - 2009 airport revenues for large hubs (Federal Aviation Administration 2010a) 29

Figure 2.3 - Process for creating model of airport stakeholders and database of their goals 30

Figure 2.4 – Organizations as customers of the airport........ 37

Figure 2.5 – Financial, Customer, and Other Relationships Between Airport Stakeholders 61

Figure 2.6 - Diagram of the economic activity positively reinforcing loop (highlighted in yellow) 67

Figure 2.7 - Diagram of the capacity and funding positively reinforcing loop (highlighted in red) 68

Figure 2.8 - Diagram of the noise and emissions negatively reinforcing loop (highlighted in green) 69

Figure 2.9 - Benchmarking Methodology Decision Tree 74

Figure 2.10 – BCC example........ 85

Figure 2.11 - Free Disposal Hull feasible region (shaded)........ 91

Figure 2.12 – Example of efficient and inefficient frontier........ 100

Figure 3.1 - Overview of airport benchmarking methodology 133

Figure 3.2 - Process for selecting metrics 142

Figure 3.3 – Overview of application of DEA model framework and heuristics 145

Figure 3.4 - Structure of a DEA model framework for airport benchmarking. Full details available in Appendix A. 146

Figure 3.5 - Architecture of the two DEA implementations 156

Figure 4.1 - Relationship between metropolitan area population and the number of domestic flights for the metro areas in Table 4.3, 2005-2008 196

Figure 4.2 – Relationship between metropolitan area GDP and the number of domestic flights for the metro areas in Table 4.3, 2005-2008 196

Figure 4.3 – Annual GDP per capita (thousands of US$), 2005-2008....................197

Figure 4.4 - Total population of metropolitan areas in millions, average 2005-2008 213

Figure 4.5 - GDP by metropolitan area in billions of US$, average 2005-2008 214

Figure 4.6 –Number of non-hub domestic destinations served nonstop, average 2005-2008 214

Figure 4.7 - Daily service frequency to top 5 hubs, average 2005-2008....................215

Figure 4.8 – Average levels of Air Service 2005-2008. 1.00 indicates the best level, and high values indicate poor service. Bar coloring is based on a k-means cluster analysis of benchmark scores. 219

Figure 4.9 - Visualization of levels of air service (Honolulu omitted), 2005-2008 average. 1.00 (dark blue) indicates the best level of air service and high values indicate poor levels of service 220

Figure 4.10 - Percentage of airline on-time arrivals in U.S. metropolitan areas; average for 2005-2008. On-time arrivals for regions with multiple airports are weighted by the volume of operations at each airport in the region. 230

Figure 4.11 – High-tech industry as a percentage of total regional GDP. Cincinnati, Denver, and Washington-Baltimore omitted due to lack of data. High-tech industry defined as NAICS categories of mining, utilities, manufacturing, information, and professional, scientific and technical services. Average based only on those years

where data in the largest number of categories was available for each city; some categories are marked in the data source as "not shown in order to avoid the disclosure of confidential information". (Bureau of Economic Analysis, U.S. Department of Commerce 2010)........238

Figure 4.12 - Level of air service as a function of high-tech industry as a percentage of total regional GDP (methodology description in caption for Figure 4.11). Marker coloring is based on a k-means cluster analysis of benchmark scores........239

Figure 4.13 – Yield by metropolitan area in relation to air service level. The best air service level is at 1.00 and the worst service level is at 3.00. Marker coloring is based on a k-means cluster analysis of the benchmark results........242

Figure 4.14 - Results from weight boundary sensitivity tests. Test 1 has the least restrictive weight boundaries, and Test 7 has the most restrictive boundaries 247

Figure 4.15 - Results from hub definition sensitivity test. The bars show the average score, and the error bars show +- the standard deviation. 250

Figure 4.16 - Number of non-hub domestic destinations served nonstop, average 2005-2008........277

Figure 4.17 - Daily service frequency to top 5 hubs, average 2005-2008........277

Figure 4.18 - Annual number of enplaned domestic passengers, average 2005-2008 278

Figure 4.19 – Capacity for flights every 15 minutes, adjusted for domestic flights only, average 2005-2008........278

Figure 4.20 - Average level of airport capacity utilization performance. 1.00 indicates strong utilization and high values indicate poor utilization. Bar coloring is based on a k-means cluster analysis of benchmark results. 281

Figure 4.21 - Capacity utilization performance in relation to average taxi-out delays. A high capacity utilization value indicates poor performance........285

Figure 4.22 - Capacity utilization in relation to gate arrival delay. A high capacity utilization value indicates poor utilization........287

Figure 4.23 - Number of annual domestic flights per airport capacity unit; average 2005-2008........289

Figure 4.24 - Average number of seats per flight (also known as gauge), 2005-2008 290

Figure 4.25 - Average load factor for departing domestic flights, 2005-2008. Although load factors can be computed in different ways, this simplified measure was computed by summing the number of enplaned passengers and dividing by the total number of seats on departing flights 292

Figure 4.26 - Regional level of air service in relation to airport capacity utilization. For both values, 1.00 indicates best performance and the highest values represent the worst performance. 294

Figure 4.27 - Level of capacity utilization in relation to the portion of domestic passengers which are O&D passengers. Average 2005-2008. A high capacity utilization value indicates poor capacity utilization. 297

Figure 4.28 - Results from weight boundary sensitivity tests. Test 1 has the least restrictive weight boundaries, and Test 7 has the most restrictive boundaries 303

Figure 4.29 - Results from hub definition sensitivity test. The bars show the average score, and the error bars show +- the standard deviation 307

Figure 4.30 - Relationship between new metric (x-axis) and traditional metric (y-axis) of capacity utilization. Colors of diamonds indicate the levels of gate arrival delay. 310

Figure 4.31 - Number of runways 324

Figure 4.32 - Number of gates 325

Figure 4.33 - Annual operating cost, average 2005-2008, million US$ 325

Figure 4.34 - Annual non-operating cost, average 2005-2008, million US$ 326

Figure 4.35 - Portion of flights arriving and departing on-time, average 2005-2008 326

Figure 4.36 - Annual aeronautical revenue, average 2005-2008, million US$ 327

Figure 4.37 - Annual non-aeronautical revenue, average 2005-2008, million US$. 327

Figure 4.38 - Annual number of enplaned passengers (international and domestic), average 2005-2008, millions 328

Figure 4.39 - Annual number of air carrier operations (takeoffs and landings), average 2005-2008, thousands 328

Figure 4.40 - Annual number of other operations (including general aviation and military flights; takeoffs and landings), average 2005-2008, thousands 329

Figure 4.41 - Results of original benchmark, average 2005-2008 330

Figure 4.42 - Benchmark results as a function of the number of enplaned passengers 332

Figure 4.43- Process for designing new study 334

Figure 4.44 - Debt service coverage ratio calculation (Pezzimenti & Macdonald 2005) 350

Figure 4.45 - Total number of enplaned passengers (millions). Including both domestic and international passengers, average 2005-2008. 361

Figure 4.46 - Total number of aircraft operations, including both takeoffs and landings, and including both air carrier and other operations. Average for 2005-2008. 362

Figure 4.47 - Runway capacity. Number of flights per 15 minutes, average 2005-2008. 362

Figure 4.48 - Total annual passenger arrival delay (millions of hours). Average 2005-2008. 363

Figure 4.49 - Results from operating efficiency benchmark. Average 2005-2008. 1.00 indicates best performance. Bar coloring is based on a k-means cluster analysis of benchmark results...364

Figure 4.50 – The operating efficiency benchmark results in relation to the number of enplaned passengers. A low efficiency score indicates strong performance 366

Figure 4.51 - Relationship between moderate and severe weather conditions and the operating efficiency benchmark results; average 2005-2008.................................369

Figure 4.52 - Regional population growth, 2008. Computed as the change from 2007 to 2008. 373

Figure 4.53 - Regional Gross Domestic Product growth. Computed as the change from 2007 to 2008. 374

Figure 4.54 - Growth in enplaned passengers, 2008. Computed as the change from 2007 to 2008. 374

Figure 4.55 - O&D passengers as portion of all passengers. 375

Figure 4.56 - Portion of the passengers that were enplaned at this airport for the carrier with the the largest share of this airport's passengers. 375

Figure 4.57 - Portion of total OEP-35 passengers (including both international and domestic) that were enplaned at this airport. 376

Figure 4.58 - Debt service coverage ratio. The formula used for computing the ratio is provided in Figure 4.44. A high debt service coverage ratio is desirable. 376

Figure 4.59 - Non-aeronautical revenue as percentage of total revenue 377

Figure 4.60 - Regional growth component benchmark results for 2008. A lower value indicates better performance. 378

Figure 4.61 – Air service component benchmark results for 2008. A lower value indicates better performance. 379

Figure 4.62 – Financial factors component benchmark results for 2008. A lower value indicates better performance. 379

Figure 4.63 - The relationship between the volume of passengers and the results of the air service benchmark. 2008. 381

Figure 4.64 - The relationship between the volume of passengers and the results of the financial factors benchmark. 2008. 382

Figure 4.65 - The relationship between the level of domestic connecting traffic and financial factors benchmark results, 2008. A low benchmark result score indicates strong performance 385

Figure 4.66 – Airport credit rating as a function of the airport's average ranking in the investment quality benchmark, 2008. 390

Figure 4.67 - Comparison of ranking of results from original benchmark, operating efficiency benchmark, and investment quality benchmark, 2008. 392

Figure 5.1 - Summary of airport benchmarking methodology 402

Figure 5.2- Summary of benchmark results. The coloring is based on the k-means clustering of each benchmark's results, ranging from dark green (best performance) to red (worst performance)..406

Figure 5.3 - Capacity utilization benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009) 410

Figure 5.4 - AIP funding directed to runway construction projects 413

Figure 5.5 - Level of air service benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009) 414

Figure 5.6 – Ranking of investment quality benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009) 416

Figure A.7 – The process for creating the new DEA model selection framework 434

LIST OF ABBREVIATIONS

ACI	Airports Council International
ACI-NA	Airports Council International – North America
AIP	Airports Improvement Program
APPS	Anti-Production Possibility Set
ATL	Hartsfield - Jackson Atlanta International
ATRS	Air Transport Research Society
AXEF	Aggressive Cross-Efficiency
BAA	British Airports Authority
BCC	Banker, Charnes, Cooper DEA model
BEA	Bureau of Economic Analysis
BOS	General Edward Lawrence Logan International
BTS	Bureau of Transportation Statistics
BUR	Bob Hope Airport
BWI	Baltimore/Washington International Thurgood Marshall
CBP	Customs and Border Protection
CCR	Cooper, Charnes, Rhodes DEA model
CLE	Cleveland-Hopkins International
CLT	Charlotte/Douglas International
CRS	Constant Returns to Scale
CVG	Cincinnati/Northern Kentucky International
DAL	Dallas Love Field
DAY	James M Cox Dayton International
DCA	Ronald Reagan Washington National
DEA	Data Envelopment Analysis
DEN	Denver International
DFW	Dallas/Fort Worth International
DMU	Decision-Making Unit
DTW	Detroit Metropolitan Wayne County
EPS	Earnings Per Share
EWR	Newark Liberty International

FAA	Federal Aviation Administration
FBO	Fixed-Base Operators
FDH	Free Disposal Hull
FLL	Fort Lauderdale/Hollywood International
GARB	General Airport Revenue Bond
GDP	Gross Domestic Product
HNL	Honolulu International
HOU	William P Hobby
IAD	Washington Dulles International
IAH	George Bush Intercontinental/Houston
IF	Inefficient Frontier
ISP	Long Island MacArthur
JFK	John F Kennedy International
LAS	McCarran International
LAX	Los Angeles International
LGA	La Guardia
LGB	Long Beach /Daugherty Field/
MCO	Orlando International
MDW	Chicago Midway International
MEM	Memphis International
MHT	Manchester
MIA	Miami International
MPO	Metropolitan Planning Organizations
MSA	Metropolitan Statistical Area
MSP	Minneapolis-St Paul International/Wold-Chamberlain
MWAA	Metropolitan Washington Airport's Authority
NAICS	North American Industry Classification System
NGO	Non-Governmental Organization
NPIAS	National Plan of Integrated Airport Systems
O&D	Origin & Destination
OAI	Office of Airline Information
OAK	Metropolitan Oakland International

OEP	Operational Evaluation Partnership
OMB	Office of Management and Budget
ONT	Ontario International
ORD	Chicago O'Hare International
PBI	Palm Beach International
PDX	Portland International
PFC	Passenger Facility Charges
PHL	Philadelphia International
PHX	Phoenix Sky Harbor International
PIT	Pittsburgh International
PVD	Theodore Francis Green State
RCP	Radius of Classification Preservation
SAN	San Diego International
SBM	Slacks-Based Measure
SEA	Seattle-Tacoma International
SFO	San Francisco International
SIC	Standard Industrial Classification
SJC	Norman Y. Mineta San Jose International
SLC	Salt Lake City International
SNA	John Wayne Airport-Orange County
STL	Lambert-St Louis International
SXEF	Simple Cross-Efficiency
TFP	Total Factor Productivity
TPA	Tampa International
TSA	Transportation Security Administration
VFP	Variable Factor Productivity
VRS	Variable Returns to Scale

ABSTRACT

A METHOD FOR STAKEHOLDER-BASED COMPARATIVE BENCHMARKING OF AIRPORTS

David Schaar, PhD

George Mason University, 2010

Dissertation Director: Dr. Lance Sherry

Major U.S. airports are critical nodes in the air transportation network, providing the interface between ground and air transportation. Airports are geographic monopolies with multiple stakeholders. Government regulations require them to operate as public utilities under profit-neutral financial conditions. By their nature, the airport stakeholders have different and sometimes conflicting performance objectives.

Since U.S. airports operate under profit-neutral regulations, enterprise performance cannot be measured using traditional financial objectives and must instead be evaluated based on the airports' ability to meet the objectives of all of

their stakeholders. Comparative benchmarking is used for evaluating the relative performance of airports.

An analysis of past benchmarks of airport performance described in this dissertation shows that these benchmarks are ambiguous about which stakeholders' needs they address and provide limited motivation for why particular performance metrics were used. Furthermore, benchmarks of airport performance use data of multiple dimensions, and such benchmarking without knowledge of utility functions requires the use of multi-objective comparison models such as Data Envelopment Analysis (DEA). Published benchmarks have used different DEA model variations with limited explanation of why the models were selected. The choices of performance metrics and the choice of DEA model have an impact on the benchmark results. The limited motivation for metrics and model render the published benchmark results inconclusive.

This dissertation describes a systematic method for airport benchmarking to address the issues described above. The method can be decomposed into three phases. The first phase is the benchmark design, in which the stakeholder goals and DEA model are selected. The selection of stakeholder goals is enabled by a model of airport stakeholders, their relationships, and their performance objectives for the airport. The DEA model is selected using a framework and heuristics for systematically making DEA model choices in an airport benchmark.

The second phase is the implementation of the benchmark, in which the benchmark data is collected and benchmark scores are computed. Benchmark scores are computed using the implementation of DEA models provided in the dissertation.

In the third phase, the results are analyzed to identify factors which contribute toward strong performance and poor performance, respectively, and to provide recommendations to decision- and policy-makers.

The benchmark method was applied in three case studies of U.S. airports:

The first case study provided a benchmark of the level of domestic passenger air service to U.S. metropolitan areas. The frequency of service to hub airports and the number of non-hub destinations served were measured in relation to the size of the regional economy and population. The results of this benchmark showed that seven of 29 metropolitan areas have the highest levels of air service. Nine areas, including Portland, OR, San Diego, and Pittsburgh, have poor levels of air service. Contributing factors to poor levels of air service are the lack of airline hub service, limited airport capacity, and low airline yields.

In the second case study, a benchmark of the degree of airport capacity utilization was conducted. The degree of capacity utilization at 35 major U.S. airports was evaluated as defined by the level of air service and volume of passengers carried in relation to the airport runway capacity. Seven out of 35 airports have the highest levels of capacity utilization while six airports, including HNL, PDX, and PIT, have poor levels of capacity utilization. Some airports with high levels of airport capacity utilization incur large delay costs while the airports with poor levels of utilization have excess capacity, indicating that funding for capacity improvements should be directed away from the poorly performing airports to those that are capacity constrained.

The third case study recreated of an existing widely published benchmark. This analysis took the premise of a previously conducted benchmark that measured airport efficiency and recreated it by applying the new benchmarking methodology in two new component benchmarks:

- A benchmark focused on the airports' operating efficiency, using parameters which included the number of passengers and aircraft movements in relation to runway capacity and delay levels
- A benchmark comparing the level of investment quality of the airports, using factors such as the debt service coverage ratio, the

portion of origin and destination passengers, and the levels of non-aeronautical revenues

The results of the new benchmark showed no statistically significant correlation with the results of the original benchmark, leading to a different set of conclusions from the new benchmarks. This illustrates the importance of a comprehensive and systematic approach to the design of a benchmark.

Practical implications of the analysis for policymakers relate to the allocation of funding for capacity improvement projects. Airports in some areas operate at high levels of capacity utilization and provide high levels of air service for their regions. These areas are at risk of not being able to satisfy continued growth in air travel demand, limiting the potential for the areas' future economic development. The most strongly affected area in this category is New York City. Similarly, the analysis found areas where the current level of air service is limited due to airport capacity constraints, including Philadelphia and San Diego. While airport capacity growth is subject to geographical and other restrictions in some of these areas, increased capacity improvement funds would provide a high return on investment in these regions.

In contrast, the analysis found that several airports with comparatively low levels of capacity utilization received funding for increased capacity in the form of new runway construction. These airports include Cleveland, Cincinnati, St. Louis, and Washington-Dulles.

In light of this indication that improvement funding is currently not optimally allocated, this benchmarking method could be used as a systematic, transparent means of enhancing the process of funding allocation.

Chapter 1: Introduction

The U.S. air transportation system is a critical component of the nation's economy, accounting for some 12 million direct and indirect jobs and $1.3 trillion in economic activity (Federal Aviation Administration 2009). However, demand is in excess of capacity in some parts of the system (Donohue et al. 2008). Estimates for the system-wide annual cost of delays range from $8 billion (Hansen et al. 2009, p. 25) to $40.7 billion (Schumer 2008).

Airports provide regional residents and businesses with access to air transportation services. Because airports require large capital investments, there is a tendency toward geographical monopolies for airports and they have been organized in a utility-like manner to prevent them from extracting excessive rates. Similar to publicly owned utilities, airports provide infrastructure to service providers and their supply chain under "revenue neutral" financial regulations (Carney & Mew 2003, p. 230). In turn, the service providers deliver seamless, safe, and secure service to the consumers of air travel services.

Research has shown that the airport exists in a form of symbiosis with its regional economy in which activity growth at the airport fuels growth in the regional economy (Button & Stough 2000), leading to regional economic growth that drives increases in the levels of activity at the airport (Intergovernmental Panel on Climate Change 2000). It has been shown that growth in Gross Domestic Product contributes about two thirds of growth in air travel, as a result of increased business activities, increased personal incomes, and increased propensity to travel (Intergovernmental Panel on Climate Change 2000).

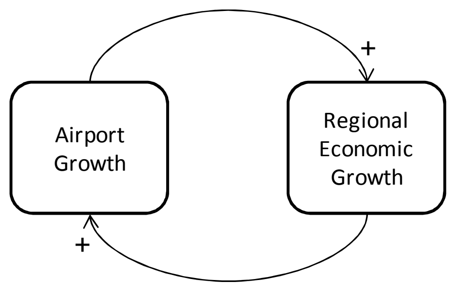

Figure 0.1 – Positively reinforcing relationship between airport growth and regional economic growth

Airport operators are charged with building the infrastructure, leasing it to the service providers, and tracking the service providers to ensure that a quality service is delivered to customers, ultimately supporting the growth of the regional economy.

A common measure of performance of for-profit enterprise performance is the earnings per share (EPS) (Tracy 2009, p. 134). Since airports function as not-for-profit utilities, such a measure is not applicable for airports. Instead, to gauge how well the nodes in the national air transportation system are achieving their performance goals, and to manage change and growth, airport operators use alternate methods. A common technique is the use of benchmarking.

A review of the published benchmarks of U.S. airports presented in section 2.2 found that:

1. The benchmarks are ambiguous to stakeholders. Many stakeholders (e.g. the local residents and business community) are ignored in these airport benchmarks, and conflicting stakeholder objectives for the airport are not acknowledged.
2. While the benchmarks all study some form of the same problem – how efficient U.S. airports are – they employ different analytical models in computing the relative performance of each airport. Research described in section 2.3.2 shows that benchmark results can be strongly impacted by the choice of analytical model.

This dissertation addresses these two gaps – 1) the lack of a stakeholder foundation of the benchmarks and 2) the lack of benchmarking model standardization – by creating an alternate methodology for stakeholder-based comparative benchmarking. Three case studies of U.S. airport performance, applying the alternate methodology, are included.

This chapter outlines the scope of the dissertation in section 1.1, describes the problems addressed by the dissertation and its unique contributions in section 1.2, provides an overview of the dissertation method in section 1.3, and gives a summary of the results in section 1.4.

1.1 Scope of the Dissertation

The scope of this dissertation is the airports that make up the Operational Evaluation Partnership (OEP) – 35 (FAA 2009). These 35 are among the U.S. airports with the highest levels of traffic, as listed in Table 0.1. This section provides the rationale for selecting this scope.

Table 0.1 - OEP-35 airports (FAA 2009)

Airport Name	Airport Code
Hartsfield - Jackson Atlanta International	ATL
General Edward Lawrence Logan International	BOS
Baltimore/Washington International Thurgood Marshall	BWI
Cleveland-Hopkins International	CLE
Charlotte/Douglas International	CLT
Cincinnati/Northern Kentucky International	CVG
Ronald Reagan Washington National	DCA

Airport Name	Airport Code
Denver International	DEN
Dallas/Fort Worth International	DFW
Detroit Metropolitan Wayne County	DTW
Newark Liberty International	EWR
Fort Lauderdale/Hollywood International	FLL
Honolulu International	HNL
Washington Dulles International	IAD
George Bush Intercontinental/Houston	IAH
John F Kennedy International	JFK
McCarran International	LAS
Los Angeles International	LAX
La Guardia	LGA
Orlando International	MCO
Chicago Midway International	MDW
Memphis International	MEM
Miami International	MIA
Minneapolis-St Paul International/Wold-Chamberlain	MSP
Chicago O'Hare International	ORD
Portland International	PDX

Airport Name	Airport Code
Philadelphia International	PHL
Phoenix Sky Harbor International	PHX
Pittsburgh International	PIT
San Diego International	SAN
Seattle-Tacoma International	SEA
San Francisco International	SFO
Salt Lake City International	SLC
Lambert-St Louis International	STL
Tampa International	TPA

In a benchmark, the comparability of the entities – airports or otherwise – is of high importance. If the entities do not have enough in common in the form of objectives, operating environment, etc., the results of the benchmark are of limited value since it cannot be determined if the better performance of entity "A" compared to entity "B" is the result of better management of entity "A" or simply due to its more favorable external environment. Accordingly, a benchmark scope must be determined that ensures good comparability.

The factors which must be considered in setting the scope for a benchmark are those that can cause poor comparability between airports if not accounted for.

By addressing these comparability issues, the quality of the benchmark results is increased. The factors which can cause poor comparability in airport benchmarking include (Mackenzie-Williams 2005):

- **Activity makeup**: Do some airports handle activities such as air traffic control while others do not, causing a different cost structure among airports?
- **Ownership**: Are some airports privately owned while others are publicly owned, causing different incentives for management?
- **Accounting practices**: Are there differences in how airports in the benchmark account for costs and revenues?
- **Funding sources**: Are some airports privately funded while others have access to government funds and grants?
- **Passenger service standards**: Are the service standards higher for some airports than for others?
- **Traffic mix**: Do some airports have mostly short-haul, domestic flights while others have more long-haul and international traffic?

Many of these comparability challenges are addressed in this dissertation by limiting the scope to the U.S. OEP-35 airports:

- **Limiting to U.S. airports only**: Limiting the scope of airports to U.S. airports ensures that the general operating environment for all airports is similar since these airports all share the same public ownership form and is subject to the same regulatory framework. Excluding airports from Europe and other parts of the world ensures that no private, for-profit airports are included in the benchmark.
- **Limiting to the OEP-35 airports**: Limiting the scope to 35 of the largest U.S. airports provides improved comparability since these larger airports share similarities in the form of operating in an environment where demand at times surpasses available capacity; smaller, secondary airports may not be subject to similar operating challenges.

The comparability challenges that are not immediately addressed by setting the scope to the OEP-35 airports are addressed individually in the case studies in section 4.

1.2 Problem Statement and Summary of Unique Contributions

1.2.1 Problem Statement

The dissertation's problem statement consists of three components:

1. **Stakeholder and goal ambiguity**: Past airport benchmarks have not examined airport stakeholders and their goals in selecting the performance metrics in use in those benchmarks. As a result, the conclusions of those benchmarks lack relevance in relation to the true goals of the airport.
2. **Lack of systematic model selection**: Past airport benchmarks lack a systematic approach to the selection of benchmarking model and many lack a motivation for why the model used was selected. As will be shown in the literature review, model selection impacts the study outcomes. The lack of systematic model selection presents the question of whether past benchmark results were valid.
3. **No benchmarks apply a systematic process**: As a consequence of problems 1 and 2 in this list, no benchmarks exist which have applied a holistic, systematic approach to benchmark design and execution.

1.2.2 Unique Contributions of Dissertation

This dissertation addresses the three problems described in section 1.2.1 through the unique contributions listed in Table 0.2.

Table 0.2 - Unique contributions of dissertation

Problem	Unique Contribution of Dissertation
1. Stakeholder and goal ambiguity	**A stakeholder model and goals database**: The dissertation presents a model of U.S. airport stakeholders and their goals for the airport in section 2.1.3.
2. Lack of systematic model selection	**Framework and heuristics for airport DEA model selection**: The dissertation presents a framework and associated heuristics for selecting a DEA model for airport benchmarking. The heuristics are based on the modeler's inputs regarding the characteristics of the airport benchmark being conducted. The framework and heuristics are presented in section 3. The framework and heuristics are applied to test the validity of past airport benchmarks in section 4.1.

Problem	Unique Contribution of Dissertation
3. Benchmarks of U.S. airports do not apply a systematic process	**Three case studies applying the new benchmark methodology**: The dissertation includes three case studies which were conducted using the new benchmark methodology. The case studies are presented in section 4 and encompass: a) A benchmark of the level of domestic air service to U.S. metropolitan areas (section 4.2) b) A benchmark of the level of capacity utilization and U.S. airports (section 4.3) c) Redesign of a past airport benchmark using the new methodology (section 4.4)

1.3 Summary of Dissertation Method

The dissertation method provides an alternative methodology for conducting airport benchmarking by identifying airport stakeholder objectives and systematically selecting a benchmarking model. A summary of the benchmarking methodology is shown in Figure 0.2.

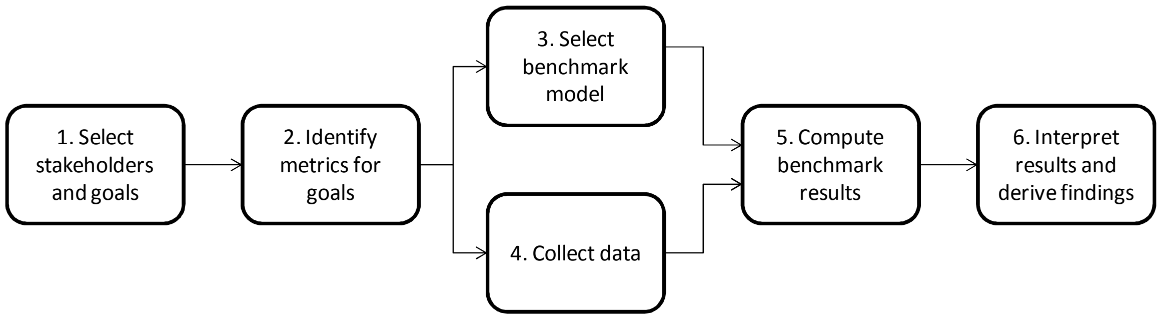

Figure 0.2 - Overview of the benchmarking methodology

The six steps of the methodology can be grouped into three phases:

1. **Benchmark design phase**: The benchmark design phase is composed of steps 1, 2, and 3 which serve to select the stakeholders and their goals to be reflected in the benchmark, to determine the appropriate performance metrics which represent those goals, and to select a DEA model which reflects the underlying characteristics of the domain being modeled. Step 1 is supported by a model of U.S. airport stakeholders and their goals for the airport which was developed as part of this dissertation. Step 3 is enabled by a framework for DEA model selection, and by heuristics for making choices in that

framework when modeling airport performance, both of which were developed as part of the dissertation.

2. **Benchmark implementation phase**: The benchmark implementation phase is made up of steps 4 and 5, in which performance data is collected and benchmark scores are computed. The computation of benchmark scores is enabled by implementations in C++ and Matlab of several DEA model variations, developed as part of the dissertation.

3. **Interpretation of benchmark results**: The final phase of the benchmark process is analyzing and interpreting the benchmark results. The purpose of this step is to uncover the factors – both controllable and uncontrollable – which impact the performance of airports in the benchmark. This provides insights to airport management, policymakers, and other stakeholders which can support decision-making related to funding, management practices, or other aspects related to the operation of airports.

1.4 Summary of Results

The analysis generated conclusions about methods for airport benchmarking as well as conclusions for policymakers about the practical implications of the results.

The analysis provided three conclusions about methods for airport benchmarking:

1. Stakeholders have different, and sometimes conflicting, objectives for the airport. A benchmark of airport performance should include a determination of which stakeholder goals are reflected, and these goals provide the basis for selecting performance metrics for the benchmark. Past benchmarks have not done so, resulting in a lack of validity of those benchmarks' results.
2. The selection of DEA model in the benchmark impacts benchmark results, sometimes to the degree of reversing the findings of the benchmark study. The selection of DEA model should take a systematic approach to ensuring that the model selected is reflective of the environment being modeled. Past airport benchmarks have not used such a systematic evaluation of models for the selection of DEA

model, which also contributes to a lack of validity of those benchmark results.

3. This dissertation provides a methodology for systematically identifying stakeholder goals and determining a DEA model whose implicit assumptions are aligned with the domain being modeled. By applying this methodology, future benchmarks can avoid the validity limitations of past benchmarks.

The benchmarking methodology was applied in three benchmarks:

1. A benchmark of the level of domestic passenger air service to U.S. metropolitan areas in relation to the size of the regional economy and population. The results showed that seven of 29 metropolitan areas have the highest levels of air service, as defined by frequency of service to hubs and the number of destinations served nonstop. Nine areas, including Portland, OR, San Diego, and Pittsburgh, have poor levels of air service. Contributing factors to poor levels of air service are the lack of airline hub service, limited airport capacity, and low airline yields.
2. A benchmark of the degree of airport capacity utilization, as defined by the level of air service and volume of passengers carried in relation to the airport runway capacity. Seven out of 35 airports have the

highest levels of capacity utilization while six airports, including HNL, PDX, and PIT, have poor levels of capacity utilization. Some airports with high levels of airport capacity utilization incur high delay costs while the airports with poor levels of utilization have excess capacity, indicating that funding for capacity improvements should be directed away from the poorly performing airports to those that are capacity constrained.

3. A benchmark that took the premise of a previously conducted benchmark from the literature and recreates it by applying the new benchmarking methodology. The results of the new benchmark diverge to a high degree from those of the original benchmark, illustrating the importance of a comprehensive and systematic approach to the design of a benchmark.

For policymakers, the practical conclusions from the analysis include:

- The finding that some areas and airports have a high need for improvement funding to improve from their current performance position, and that other areas and airports do not have that need. Areas such as Philadelphia and San Diego are reporting levels of air service that is not in line with the best-in-class and areas such as New York City are at the risk of proportionally reduced levels of air service in the future, and this is due to airport capacity limitations. In contrast, airports such as Cleveland, Cincinnati, St. Louis, and Washington-Dulles are receiving improvement funding for runway construction in spite of low relative levels of capacity utilization. This suggests that the method for airport benchmarking presented in this dissertation could be used for optimizing the allocation of improvement funding.

- A caution about the use of benchmarking to support decision-making if there is no transparency about the stakeholder goals that are reflected in the benchmark results and if no explanation is provided about how the DEA model used is reflective of the domain being modeled. If there is no clear understanding about these factors, it cannot be determined whether the benchmark results are actually indicative of strong and poor airport performance, as defined by the airport goals.
- The insight that airports serve multiple stakeholders with different and sometimes conflicting interests. The airport functions in a utility-like role in its region and its goals are to meet the objectives of these stakeholders. The model of stakeholders and their goals for the airport provided in the dissertation identifies the incentives and

motivations for the behavior of airport management and other stakeholders in the airport system.

2 Chapter 2: A Review of Airport Benchmarking

The literature review consists of a discussion of the airport's role as a public utility with multiple stakeholders in the first subsection, a review of airport benchmarking and analytical techniques for benchmarking in the second subsection, and a review of the dissertation's problem statement in the third subsection.

2.1 The Airport as a Public Utility

This section discusses the concept of public utilities, reviews airport financing, and reviews airport stakeholders and their goals for the airport.

2.1.1 Public Utilities

Major U.S. airports function as public utilities. Utilities (e.g. electric distribution utilities) require high capital investments for system construction. Duplication of system infrastructure is considered inefficient, and as a result utilities operate in some monopolistic form (White 1976, p. 14). For instance, the definition of an electric utility makes the distinction that it is a monopoly: An electric utility is "[a]ny organization, municipality or state agency with a monopoly franchise that sells electric energy to end-use customers" (Public Utility Research Center, University of Florida 2008).

Utility ownership is either public (federal, state, or municipal) or private (Schap 1986, p. 3). In the cases of private ownership, strict regulation is in place to ensure that the monopolistic situation is not used to charge excessive prices (Hunt 1944, pp. 16-17). Utility regulation exists "to assure to ultimate consumers the best possible service at reasonable cost" (Hunt 1944, p. 17). For example, quality electricity distribution service is defined as "the uninterrupted flow of current and [...] the ability to maintain constant frequency voltage within the limits that will ensure satisfactory performance of the consumer's equipment and appliances" (White 1976, p. 9).

Public utilities have a number of different stakeholders, including shareholders/creditors (if applicable), government regulators, and customers. Given this operating situation, a utility's performance of its mission cannot be gauged only by its ability to generate profits. Instead, the interests and considerations of all of the utility's stakeholders must be considered in evaluating the utility's performance, in particular in the cases where utilities are under some form of government ownership. This can be a complex problem since stakeholders may have conflicting objectives.

Similarly, airports exist to provide a quality service to regional businesses and residents at a reasonable price, while generally operating in a monopolistic (or semi-monopolistic) environment. All major airports in the United States are publicly owned enterprises financed by a combination of public and private funds, and are barred from generating a financial surplus (Carney & Mew 2003, p. 230). Rather than comparing profitability, airports' performance must, similar to other public utilities, be gauged by their ability to meet the interests of all of its stakeholders.

2.1.2 Airport Finance

Airports are dependent on capital funding for infrastructure development and on revenues for covering the costs of operations, depreciation, and interest. Understanding airport financing is a necessary basis for an examination of airport stakeholders since groups of stakeholders with sometimes conflicting objectives contribute toward the funding of airports. This section discusses sources and types of capital funds, and sources of airport revenues.

2.1.2.1 Airport Capital Funding

Airports require access to sources of capital funding for infrastructure development projects. Projects such as runway additions, terminal expansion projects, and purchase of capital-intensive equipment (e.g. fire trucks) are considered capital improvement expenses (Wells & Young 2003, p. 311). In their role as public utility-like entities, airports interact with several different stakeholders that provide capital funding. Five key sources of capital funding exist for the airport:

- FAA Airports Improvement Program (AIP) (G. Hamzaee & Vasigh 2000)
- Bonds (G. Hamzaee & Vasigh 2000)
- Airport operating surplus (G. Hamzaee & Vasigh 2000)
- Passenger Facility Charges (PFCs) (G. Hamzaee & Vasigh 2000)
- State and local funding (Airports Council International - North America 2009a, p. 22)

Figure 2.1 shows the average capital funding source breakouts for large hubs[1].

[1] Large hubs are defined as having at least 1% of total annual passenger boardings (Airports Council International - North America 2009a, p. 22)

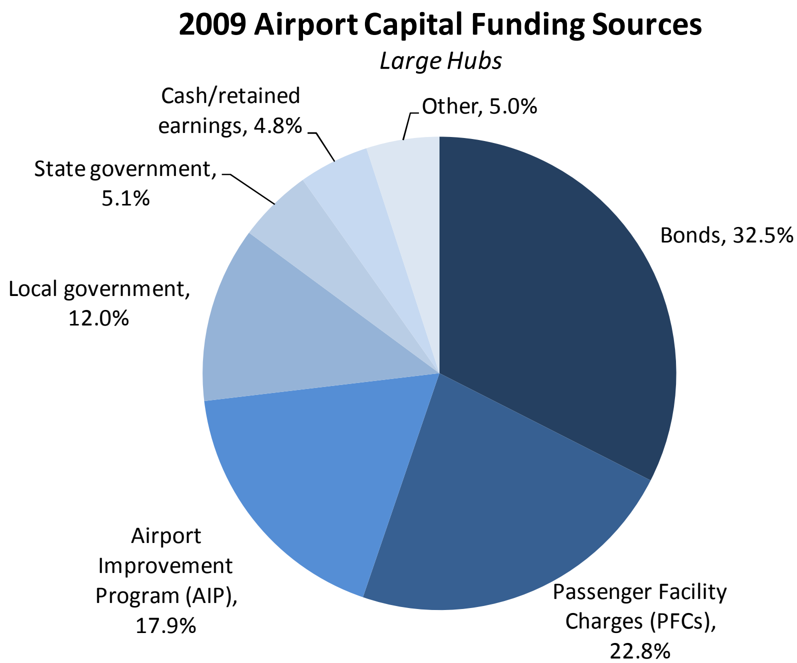

Figure 2.1 - Airport capital funding sources for large hubs (Airports Council International - North America 2009a, p. 10)

2.1.2.2 Airport Revenues

Airport revenues come from different sources and are categorized as follows (Federal Aviation Administration 2001):

- **Aeronautical operating revenue**: Including landing fees, terminal rental fees, apron charges, FBO revenue, cargo and hangar rentals, and aviation fuel taxes.

- **Non-aeronautical operating revenue**: Including terminal revenue (including food and beverage and retail revenue), rental car revenue, and parking revenue.
- **Nonoperating revenue**: Interest income, grants, and Passenger Facility Charges

The largest source of revenues for large hubs is aeronautical revenue, as shown in Figure 2.2.

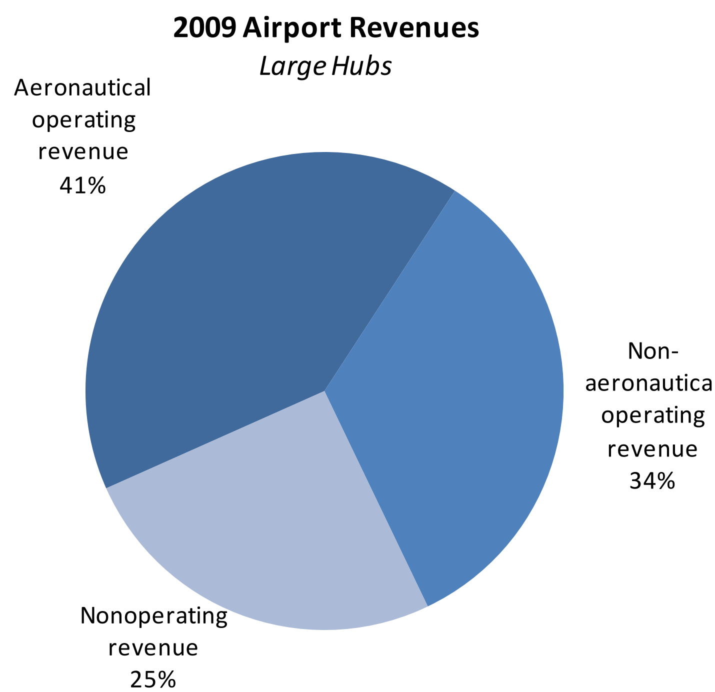

Figure 2.2 - 2009 airport revenues for large hubs (Federal Aviation Administration 2010a)

2.1.3 A Model of U.S. Airport Stakeholders and their Goals for the Airport

The purpose of this section is to present a model of the stakeholders in U.S. airports and to analyze the stakeholders' goals for the airport. This section includes an identification of who the stakeholders are, an analysis of their definitions and goals, and then presents a model of the stakeholder relationships.

A model of stakeholders in U.S. airports and their goals did not previously exist. This new model was created through the process described in Figure 2.3.

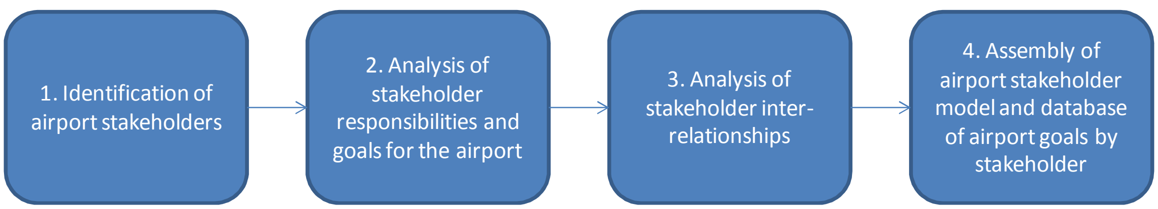

Figure 2.3 - Process for creating model of airport stakeholders and database of their goals

The model construction relied on two categories of sources:

1. A review of literature on airport management, finance, legislation, and other topics.
2. Knowledge elicitation sessions with representatives of airport stakeholder groups. These sessions were conducted during 2009

both in-person and via phone. 32 sessions were conducted in all and they consisted of questions regarding stakeholder definitions, goals of those stakeholders, and key performance indicators relevant to these goals. The list of subject-matter experts, their titles, and affiliations and the knowledge elicitation form are presented in Appendix D.

2.1.3.1 Identification of Airport Stakeholders

With the airport operating as a public utility, an inventory of airport stakeholders and their objectives is required to form the basis for evaluating the airport's performance.

For the purpose of this analysis a stakeholder is defined as "any group or individual who can affect or is affected by the achievement of the organization's objectives" (Mitchell et al. 1997, p. 856). Table 2.1 describes a comprehensive list of stakeholders generated through a review of the literature.

Table 2.1 - Airport stakeholders

Stakeholder Group	References Citing Group
Passengers	(Upham 2003) (Rhoades et al. 2000) (Neufville & Odoni 2003)
Air carriers	(Upham 2003) (Rhoades et al. 2000) (Offerman) (Neufville & Odoni 2003) (Sarkis & Talluri 2004)

Stakeholder Group	References Citing Group
General aviation users	(Rhoades et al. 2000)
Airport organization	(Upham 2003) (Rhoades et al. 2000) (Offerman) (Sarkis & Talluri 2004)
Investors and bond-holders	(Neufville & Odoni 2003)
Concessionaires	(Rhoades et al. 2000) (Neufville & Odoni 2003)
Service providers	(Upham 2003) (Rhoades et al. 2000) (Neufville & Odoni 2003)
Employees	(Upham 2003)
Federal government	(Upham 2003) (Offerman) (Neufville & Odoni 2003) (Sarkis & Talluri 2004)
Local government	(Upham 2003) (Offerman) (Neufville & Odoni 2003) (Sarkis & Talluri 2004)
Communities affected by airport operations	(Upham 2003) (Offerman)
NGOs, such as environmental bodies	(Upham 2003)
Business, commerce, tourism, arts, sports, and education organizations	(Upham 2003)
Parking operators and ground transportation providers	(Upham 2003) (Neufville & Odoni 2003)
Airport suppliers	(Upham 2003) (Neufville & Odoni 2003)

2.1.3.2 Analysis of Stakeholder Definitions and Goals

To examine the role of the airport stakeholders a precise definition of stakeholders and their goals for the airport is necessary. The purpose of this section is to identify the airport's goals from the point of view of each stakeholder group.

2.1.3.2.1 Passengers

For passengers, the airport provides a transition point between the ground and air transportation modes, or a connection point between two flights. Different sub-types of passengers have been identified (Neufville & Odoni 2003, pp. 610 - 611):

- Arriving passengers
- Originating passengers
- Transfer passengers
- International and domestic passengers
- Charter and low-fare airline passengers
- Shuttle/commuter passengers

These passenger types are not mutually exclusive; rather, an individual passenger may be a member of more than one sub-type of passenger categories. Arriving and originating passengers are commonly referred to as origin and destination (O&D) passengers.

Independent of the passenger classifications according to the above attributes, the passengers may be viewed in two different capacities in the context of this analysis. First, passengers can be viewed as participants in the economic system, either as business travelers or as tourist/leisure travelers, purchasing services from airport service providers and interacting in different ways with local businesses and the local community. Second, passengers can be viewed as individual travelers that have expectations about receiving quality services, and passing through the airport system in a convenient manner. These two perspectives have different implications on the goals for the airports and will be treated separately in the following subsections.

2.1.3.2.1.1 Passengers as Economic Participants

Passengers may participate in the economic system in one of several ways:

- As origin leisure/personal travelers: These are passengers from the local community that use the airport as their departure point for leisure or other personal travel.
- As origin business travelers: These are travelers representing local businesses, using the airport as their departure point.
- As destination leisure/personal travelers: These are visitors to the region, for tourism or other personal purposes.

- As destination business travelers: These are business travelers coming to visit local businesses.

Each type of passenger has a different impact on the local region, as will be discussed in section 2.1.3.2.2.

If the airport's traffic is heavily geared toward O&D traffic, then demand at the airport is more heavily dictated by the local economy. In contrast, high connecting (transfer) passenger levels are less sensitive to the performance of the local economy, but those traffic volumes may represent a vulnerability for the airport since they are to a greater degree dictated by a carrier's viability and route decisions (Forsgren 2007, p. 2).

Passengers contribute toward the financing of airport capital improvement projects through Passenger Facility Charges (PFCs) of up to $4.50 per passenger. PFCs are paid directly by passengers through airline tickets and proceeds must be used for capital improvements at the airport that collected them (Wells & Young 2003, p. 79).

The goals for passengers as economic participants relates to the cost of travel: Providing access to low airfares is a key objective for the airport in the view of air passengers (Michael Cintron, International Air Passengers' Association 2009).

The role of passengers in the economic system is further discussed in sections 2.1.3.2.2 and 2.1.3.2.11.

2.1.3.2.1.2 Passenger as Travelers

When considering the passengers as travelers as a stakeholder group, the focus is on the passenger as an individual. The goal of the airport from the individual passenger viewpoint is "moving passengers quickly and conveniently to where they need to go" (Michael Cintron, International Air Passengers' Association 2009). This view considers the airport as a transit point from one mode of transportation to another, or as a connection point between two different flights. Ensuring on-time performance was raised as the most important aspect to achieving this objective.

2.1.3.2.2 Business, Commerce, Tourism, Arts, Sports, and Education Organizations

The organizations that in various ways are customers of the airport have been summarized as "business, commerce, tourism, arts, sports, and education organizations" (Upham 2003). Figure 2.4 proposes a means for categorizing these organizations based on the type of use they derive from the airport: Some organizations are direct users of the airport by importing or exporting services (i.e.

business travelers) and goods (raw materials or finished goods). Other organizations are indirect customers of the airport as a result of their customers (e.g. tourists) traveling through the airport.

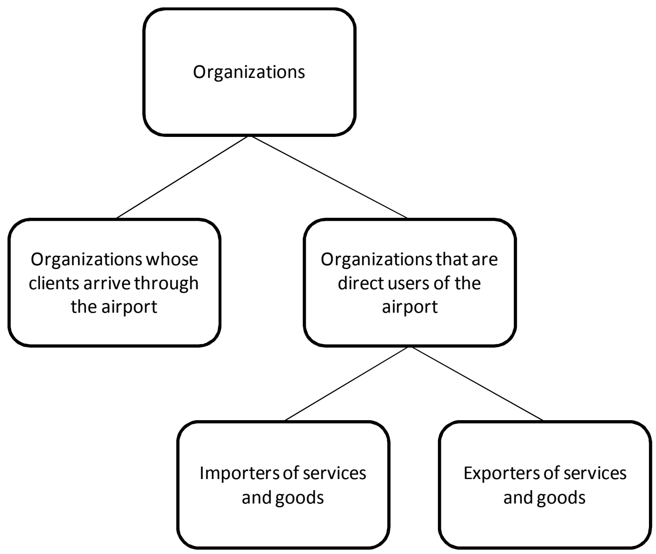

Figure 2.4 – Organizations as customers of the airport

The airport serves as an engine of business activity for the organizations in the region. The airport drives and supports economic activity in several different ways, including both through business activities directly at the airport and through business activities throughout the regional economy (Button & Stough 2000). These types of economic activity are described in greater detail in section 2.1.3.2.11. Underlying goals for maximizing this economic activity include maximizing passenger volumes and traffic at the airport as well as maximizing the number of destinations served and the frequency of those services (Matt Erskine, Greater

Washington Board of Trade 2009). As a result of the different types of use of the airport described in the previous paragraph, the priority of one goal over another varies between organizations.

2.1.3.2.3 Air Carriers

Air carriers provide the air transportation service from the airports. Air carriers include both passenger and cargo carriers and are classified into three subcategories, according to (Environmental Protection Agency 2000, p. 14-26):

- Large certified carriers: These carriers have a certificate to carry 61 passengers or more, payload equal to or greater than 18,000 pounds, or conduct international operations
- Small certified carriers: These carriers fly aircraft that carry less than 61 passengers, carry less than 18,000 pounds, and do not conduct international operations.
- Commuter carriers: These are air taxis with a published schedule of at least five weekly round trips between at least two airports.

Air carriers select airports based on the passenger demand for service to/from the airports (i.e. revenue generation potential) and based on the cost of operating at the airport. The airlines have the objective of achieving high yields, (Doganis 2002, p. 16). Airports serve the role of providing access to high yield markets. Attractive airports ensure low cost of air carrier operations at the airport. This includes both minimizing direct fees charged to air carriers through the maximization of non-aeronautical revenues (Dallas Dawson, Tampa International Airport 2009) and minimizing costs incurred by air carriers through delay on the ground (Peter Stettler, Ricondo and Associates 2009).

An airport may serve either as a hub for a carrier, with a high portion of that carrier's flights operating to/from the airport, or as a non-hub airport with a lower portion of flights for a given carrier (Belobaba et al. 2009, pp. 162-163). In either situation, the airport should act as an efficient hub/connection point, contributing to ensuring air carriers' on-time performance (Pat Oldfield, United Airlines 2009).

In addition, it is the expectation of air carriers that airports ensure safety of operations on the airport surface (Kurt Krummenacker, Moody's 2009).

2.1.3.2.4 General Aviation Users

General aviation encompasses many types of aviation outside the air carrier definition, including (Wells & Young 2003, p. 386):

- Air taxi operators (except those air taxi operators listed in section 2.1.3.2.3)

- Corporate-executive transportation
- Flight instruction
- Aircraft rental
- Aerial application
- Aerial observation
- Business
- Pleasure

Several of the goals listed for air carriers in section 2.1.3.2.3 also apply to general aviation in terms of on-time performance, low costs, and safety. However, a representative of a business aviation organization defined the primary goal of airports as serving as access point to the national air transportation system by providing good availability and high capabilities in terms of instrumentation and services (Jeff Gilley, National Business Aviation Association 2009).

2.1.3.2.5 Airport Organization

The airport organizational structure varies (Neufville & Odoni 2003, p. 225) and can be comprised of an individual airport such as Dallas Forth Worth Airport (DFW) (DFW Airport 2009) or as a group of airports managed by the same organization, such as the Metropolitan Washington Airports Authority (MWAA) (Metropolitan Washington Airports Authority 2009). The airport organization is overseen by a board appointed by local governments, as described in section 2.1.3.2.11.

In larger airports or systems of airports, a common feature is that the organization includes a separation of operating units which carry out on-going management of airport operations, and they are separate from staff units which have responsibility for (among several other areas) infrastructure development (Neufville & Odoni 2003, pp. 226-227).

The airport itself pays for some capital infrastructure projects, as shown in section 2.1.2. Airport operating revenues come from sources such as landing fees, terminal leases and proceeds from concessions sales. This revenue is used to pay for the airport's operating expense, but any surplus can be used to contribute toward capital improvements (Dillingham 1996, p. 9).

A set of goals for the airport organization can be derived from studying airports' strategic plans and objectives and from knowledge elicitation sessions with airport management experts.

The primary objective (sometimes referred to as the "mission") of the airport is to provide access to high quality air services to its region. Other goals, such as ensuring strong financial performance and high operational efficiency, are considered as "means to an end" in that they enable the airport to achieve this overarching goal (DFW Airport 2008, p. 2) (Hillsborough County Aviation Authority 2006, p. 5) (Jim Wilding, formerly with MWAA 2009).

A summary view of the airport's goals is presented using the structure of Denver International Airport's strategic plan (Denver International Airport 2009):

2.1.3.2.5.1 Goal 1: Excel in Airport Management

The goal of excelling in airport management includes:

1. Achieve high security and safety (City of Cleveland, Department of Port Control 2007, p. 6) (Denver International Airport 2009, p.8) (Hillsborough County Aviation Authority 2006, p. 5)
2. Grow revenue and manage costs (City of Cleveland, Department of Port Control 2007, p. 14) (Denver International Airport 2009, p.8) (DFW Airport 2008, p. 3) (Hillsborough County Aviation Authority 2006, p. 5)
3. Drive economic growth (Denver International Airport 2009, p.8)
4. Grow passenger numbers (City of Cleveland, Department of Port Control 2007, p. 14) (Denver International Airport 2009, p.8)
5. Provide access to a high number of destinations and a high frequency of service (Denver International Airport 2009, p.8). This goal relates immediately to the primary objective of the airport described above.

Airport management must also achieve a balance where sufficient infrastructure capacity exists for handling traffic while capacity is at the same time not over-built (Paul McKnight, Jacobs Consultancy 2009) (Frank Berardino, GRA Inc 2009). Additionally, a key objective for airports is to maximize non-aeronautical revenues since that provides diversified revenues and allows for keeping usage charges to air carriers low, thereby potentially attracting more traffic (Chellie Cameron, MWAA 2009) (Peter Stettler, Ricondo and Associates 2009) (Seth Lehman and Emma Walker, Fitch Ratings 2009).

2.1.3.2.5.2 Goal 2: Provide High Levels of Customer Service:

The goal of high levels of customer service includes ensuring a good experience for both passengers and other customers (City of Cleveland, Department of Port Control 2007, p. 7) (Denver International Airport 2009, p. 9) (DFW Airport 2008, p. 3) (Hillsborough County Aviation Authority 2006, p. 5).

2.1.3.2.5.3 Goal 3: Develop Environmentally Sustainable Practices and Minimize Noise

This goal includes minimizing emissions, energy consumption, etc., within the airport (Denver International Airport 2009, p. 10) (City of Cleveland, Department of Port Control 2007, p. 14). Some airports, such as Sea-Tac, are also beginning to expand their focus by considering greenhouse gas emissions not only from the airport-controlled operations but also from airlines and other tenants as well as the public (Port of Seattle, Sea-Tac Airport 2007, p. ES1). Related to this is also the goal of minimizing airport-related noise (Neufville & Odoni 2003, pp. 167-170).

2.1.3.2.5.4 Goal 4: Develop High-Performing Employee Teams

This goal relates to developing effective and skilled employees (City of Cleveland, Department of Port Control 2007, pp. 5, 15) (Denver International Airport 2009, p. 12) and maximizing employee engagement (DFW Airport 2008, p. 3).

2.1.3.2.5.5 Goal 5: Enhance Competitive Advantage

This goal includes providing competitive user rates and protecting the airport's physical infrastructure (Denver International Airport 2009, p. 14) (City of Cleveland, Department of Port Control 2007, p. 13).

2.1.3.2.5.6 Conflicts Between Airport Organization's Goals

Some of these goals may be in competition with each other. For instance, the goal of maximizing non-aeronautical revenue can conflict with the goal of developing environmentally sustainability and providing a good experience for passengers: The latter two goals would be aided by promoting and developing access to public transportation access modes to the airport such as bus or rail.

However, the goal of maximizing non-aeronautical revenue is better served by maximizing revenue-generation in the form of parking revenue from private vehicles. In such instances, airport management must balance the competing priorities in order to accomplish the goals of the airport.

2.1.3.2.6 Investors and Bond-Holders

The majority of airport debt is of the general airport revenue bond (GARB) type. GARB means that the bond is backed by revenues generated from airport operations and not backed by any government funding source. The credit ratings agencies Moody's, Standard and Poor's, and Fitch Ratings participate in this system by assigning grades of investment quality to the airports' bonds. The ratings agencies' ratings affect the interest rates and terms of the bonds (Wells & Young 2003, pp. 336-339). A large number of factors impact the bond ratings, including (Forsgren 2007, p. 2):

- Historical and projected population growth
- Historical and projected employment expansion and mix
- Passenger growth
- Airport utilization trends
- Portion of origin and destination (O&D) traffic
- The importance of the facility to the overall U.S. system of airports

- Whether the airport is in a favorable geographic location (e.g. is it a natural hub location?)
- Airfield capacity and attractiveness of facilities
- Debt burden and carrying costs
- Financial strength of carriers with a lot of connecting traffic, and their level of commitment to the airport
- The role of the airport in the dominant carrier's network
- The level of legal flexibility for the airport to change the rates it charges air carriers

2.1.3.2.7 Concessionaires

Airport concessionaires operate passenger services in terminal buildings and may include food and beverage services, retail services, and hotels. Concessions operators pay the airport organization a fixed annual fee and/or a percentage of gross revenues (Wells & Young 2003, p. 324). Considering the concessions operators' objective of maximizing profits, the goals of the airport for these operators are deduced to be maximizing passenger volumes and minimizing the fees paid to the airport organization.

2.1.3.2.8 Service Providers

The service providers are private operators that offer services to air carriers and general aviation users. Independent operators may supply these services (e.g. fixed-base operators, FBOs), but some of the services may also be provided by the airport operator, the airline itself, or by another airline. Services provided include (Neufville & Odoni 2003, pp. 268, 278):

- Supply of aviation fuel and oil
- Baggage handling and sorting
- Loading and unloading of aircraft
- Interior cleaning of aircraft
- Toilet and water service
- Passenger transport to/from remote stands
- Catering transport
- Routine inspection and maintenance of aircraft at the stands
- Aircraft starting, marshalling, and parking
- Aircraft de-icing
- Passenger handling (e.g. ticketing and check-in)
- Cargo and mail handling
- Information services
- Preparation of handling and load-control documents

- Supervisory or administrative duties

Similar to concessionaires, independent service providers pay a fee to the airport organization which is typically a percentage of gross revenues (Neufville & Odoni 2003, pp. 268, 279). In a parallel to concessionaires, service provider goals for the airport would include maximizing traffic volumes and minimizing the fees paid to the airport organization.

2.1.3.2.9 Employees

The employee category includes both direct employees of the airports organization as well as employees of companies operating at the airport, such as concessions operators. Some employees are organized into unions, such as the Service Employees International Union (SEIU USW West 2009) and Unite Here (Unite Here 2009). The objective of the airport from the perspective of those unions is to provide secure jobs, wages, and benefits (Unite Here 2009).

2.1.3.2.10 Federal Government

The federal government participates in the airport system in three different roles: As a bill-payer, as an operator, and as a regulator. Each of these roles will be addressed in this section.

In terms of the government's role as a bill payer for the system, the Airports Improvement Program (AIP) is administered by the FAA and its funding comes from the Airport and Airway Trust Fund, which in turn is funded by user fees and fuel taxes. AIP funds can be applied toward projects that "support aircraft operations including runways, taxiways, aprons, noise abatement, land purchase, and safety, emergency or snow removal equipment" (Kirk 2003, p. 3). In order to be eligible for AIP funding, airports must be part of the National Plan of Integrated Airport Systems (NPIAS), which imposes requirements on the airport for legal and financial compliance (Wells & Young 2003, p. 329).

The NPIAS has two goals: To ensure that airports are able to accommodate the growth in travel, and to keep airports up to standards for the aircraft that use them (FAA 2008, p. v).

The government's role as airport operators includes three different agencies:

- **FAA**: The FAA is the operator of ramp, ground, local, and departure/arrival air traffic control services (United States Code of Federal Regulations 2010).
- **Transportation Security Administration (TSA):** The TSA provides passenger and baggage security screening services. The TSA states that it is the goal for its baggage screening operations to screen for

explosives and other dangerous items while maximizing efficiency (Transportation Security Administration 2009). This can be translated to state that it is the goal for the airport to ensure secure transportation of people and goods while minimizing the impact of security measures on legitimate travelers and goods.

- **Customs and Border Protection (CBP):** The CBP is responsible for operating passport control and customs inspections at international airports. The CBP states that it is its mission to protect "our nation's borders from terrorism, human and drug smuggling, illegal migration, and agricultural pests while simultaneously facilitating the flow of legitimate travel and trade" (Customs and Border Protection 2009). Just as for the TSA, this can be translated to state that it is the goal for the airport to ensure secure transportation of people and goods while minimizing the impact of security measures on legitimate travelers and goods.

Lastly, the federal government is a regulator of the airports system. Airports that are included in the NPIAS are subject to a number of federal regulations that are enforced by the FAA and the TSA. The regulations apply to both the airport infrastructure as well as to service providers within the airport systems. The purpose of these rules is to ensure the safe and efficient operations of public-use airports (Wells & Young 2003, pp. 19-22).

2.1.3.2.11 Local Government

U.S. airports are with few exceptions not private, profit-making enterprises. Instead, airports are typically owned and operated by public entities such as cities, counties, or local airport authorities (Neufville & Odoni 2003).

For instance, Washington's Dulles and National airports are owned and operated by the Metropolitan Washington Airport's Authority (MWAA). The MWAA is officially a body independent of the local government but its board is appointed by the Governor of Virginia, the Mayor of the District of Columbia, the Governor of Maryland and the President of the United States).

Similarly, Newark, LaGuardia, JFK, Stewart International, and Teterboro airports in metropolitan New York City are owned by the Port Authority of New York and New Jersey (The Port Authority of New York and New Jersey 2009). Dallas-Fort Worth Airport is jointly owned by the City of Dallas and the City of Forth Worth (DFW Airport 2009).

The government owners in the form of city and local governments are represented by an airport board which is responsible for the strategic direction of the airport and for appointing airport management (Wells & Young 2003, p. 35).

The local government is supported in an advisory role by federally funded Metropolitan Planning Organizations (MPOs) who are charged with assisting in planning for aviation and other transportation infrastructure for the local region (Association of Metropolitan Planning Organizations 2010).

State and local government also contribute as bill-payers for capital improvement projects (Airports Council International - North America 2009a).

The objectives of the airport from the point of view of the local government is representative of those of the local community it represents and involves both maximizing its positive effects while minimizing its negative effects as described in the subsequent paragraphs.

One form of positive impact of the airport is in the shape of economic effects. However, many studies of the economic impact of airports are sponsored by the airports authorities themselves, making them "more political than analytical" (R Cooper 1990). Although there may be no definitive measure of the economic impact of airports, a structure for the types of impacts of airports to their regional communities has proposed (Button & Stough 2000):

1. Short-term impact from construction, expansion, and renovation of airports

2. Sustained impact in the form of jobs at the airport (direct impact) and off-airport jobs that result from the "multiplier effect" of the income generated by employees at the airport
3. Stimulus of the local economy as a result of firms and individuals having air transportation services at their disposal
4. Spurring other economic development by crossing thresholds for economies of scale, scope, and density. The authors note that this last form of impact is very difficult to quantify.

Related to the objective of maximizing economic effects is providing maximum access to air services that connect the region to the country and the world. This involves maximizing the number of destinations served and the frequency of those services (Jim Wilding, formerly with MWAA 2009) (Kurt Krummenacker, Moody's 2009) (Chellie Cameron, MWAA 2009) (Matt Erskine, Greater Washington Board of Trade 2009).

As described for airport management in section 2.1.3.2.5, the objective of the local government is also to minimize the negative impact of the airport in the form of noise and emissions.

2.1.3.2.12 Communities Affected by Airport Operations

The interest of communities affected by airport operations is represented by the local government which was elected by the constituents of those communities. Hence, the goals of the airport for these communities are broadly aligned with the goals described for the local government in the preceding section, including maximizing economic impact, maximizing destinations served and frequency, and minimizing emissions and noise.

However, it should be noted that for individual groups of community members, the objectives of the airport may be different for others. According to Smith (Smith 1979, p. 47), "how much people suffer from this growing nuisance depends largely on where they live, which may have no bearing on how much they benefit from the airport." From this reasoning, residents near the airport can be considered a particularly important subset of the overall group of communities affected by airport operations.

The adverse effects of airports result from several sources, including air traffic, ground vehicles on the airport, and vehicles providing ground transportation to travelers (Wells & Young 2003, pp. 354-361). The adverse effects include:

- Noise
- Air quality

- Water quality
- Hazardous waste emissions
- Other externalities, including increased automobile traffic congestion

2.1.3.2.13 NGOs, such as Environmental Bodies

Non-governmental organizations (NGOs), such as environmental bodies, fall in the category of "airport interest groups". Although they state that "there are many national organizational and regional organizations that are deeply interested in the operation of airports", Wells and Young (Wells & Young 2003, pp. 22-24) only list NGOs that can be considered "pro-aviation", such as the Aerospace Industries Association, the Airports Council International – North America, and the International Air Transportation Association.

However, interest groups with other interests also exist, such as environmental bodies (Upham 2003). The US-Citizens Aviation Watch is such an organization, which is "dedicated to protecting the health, safety and welfare of individuals and communities that are affected by the air transport industry" (US-Citizens Aviation Watch 2009).

What is clear from this, however, is that there can be no general description of the goal of airports representing all NGOs.

2.1.3.2.14 Parking Operators and Ground Transportation Providers

Ground transportation providers include rail service, taxicabs, buses, shuttles, rental cars, and limousines, while parking services may be provided both on and off the airport, and either by the airports organization or by private enterprises. From airport management's point of view, the desirable distribution between different modes of transportation will vary dependent upon the individual airport's context (Wells & Young 2003, pp. 229-241).

Similar to concessionaires and airport service providers, the revenues for parking operators and ground transportation providers will be maximized through high passenger volumes and (where applicable) low fees paid to the airport.

2.1.3.2.15 Airport Suppliers

Airport suppliers have the airport itself as the end customer. These include for instance various contractor and consulting firms and equipment suppliers (Upham 2003). Similar to concessions, airport service providers, and ground transportation providers, these suppliers benefit from growth in traffic volumes.

2.1.3.3 Summary of Stakeholder Definitions and Goals

The discussion in section 2.1.3.2 of stakeholders, definitions, and their goals for the airport is summarized in Table 2.2.

Table 2.2 – Description of airport stakeholders and goals

Stakeholder Group	Definition	The Stakeholder 's Goals for the Airport
Passengers	O&D and transferring passengers	• Move passengers quickly and conveniently • Ensure on-time performance • Provide access to low fares
Organizations	Organizations in region	• Maximize passenger and traffic volumes • Maximize number of destinations served and frequency of those services
Air carriers	Passenger and cargo carriers	• Ensure on-time performance • Ensure low cost of operations • Ensure safety of operations • Provide access to high yields
General aviation	Air taxi, corporate transportation, business aviation, etc.	• Serve as access point to the NAS through good availability and high equipment capability

Stakeholder Group	Definition	The Stakeholder 's Goals for the Airport
Airport organization	Individual airports or multi-airport systems, including management and staff, with responsibility for building and operating the airport	• Achieve high security and safety • Grow revenue and manage costs • Drive economic growth • Grow passenger numbers • Find opportunities for new destinations and increase service frequency • Ensure sufficient (but not excessive) infrastructure capacity • Maximize non-aeronautical revenues • Maximize customer satisfaction • Achieve environmental sustainability • Minimize noise • Develop employees • Enhance competitive advantage
Investors and bond-holders	Individuals/organizations holding bonds, and the credit ratings agencies	• Optimize performance in factors under consideration (see section 2.1.3.2.6)
Concessionaires	Operators of passenger services such as food and beverage and retail	• Maximize passenger volumes • Minimize fees paid
Service providers	Providers of services to air carriers, such as fuel	• Maximize traffic volumes • Minimize fees paid
Employees	Employees of the airport organization and airport tenants	• Provide secure jobs, wages, and benefits

Stakeholder Group	Definition	The Stakeholder 's Goals for the Airport
Federal government	Bill-payer for infrastructure (AIP), operator of air traffic control and security, and system regulator.	• Ensure that airports can accommodate growth • Keep airports up to standards • Ensure safety, security, and efficiency of operations
Local government	Local entities such as counties or cities which own airports.	• Maximize economic impact • Maximize number of destinations served and frequency of those services • Minimize noise and emissions
Communities affected by airport operations	Residents in region, and in particular residents near the airport	• Maximize economic impact • Maximize number of destinations served and frequency of those services • Minimize noise and emissions
NGOs, such as environmental bodies	Airport interest groups	• Varies depending on the interest group
Parking operators and ground transportation providers	Rail service, taxicabs, buses, shuttles, rental cars, limousines, and on and off airport parking services	• Maximize passenger volumes • Minimize fees paid
Airport suppliers	Suppliers of contractor and consulting services and equipment	• Maximize traffic volumes

2.1.3.4 A Model of Airport Stakeholder Relationships

Using the knowledge from section 2.1.3.1, a diagram of the airport stakeholders and their interrelationships can be constructed based on the descriptions of the responsibilities/needs of each stakeholder and their primary points of interactions. This section describes such a model.

2.1.3.4.1 Airport Stakeholder Model Overview

The stakeholder model is shown in Figure 2.5. At the center of the diagram are the airport organization and the physical airport infrastructure. The diagram shows that the service providers are the primary entities that interact with the airport infrastructure and that the end users in the form of passengers and cargo forwarders interact with the service providers.

Within the regional economy and local community, some stakeholders are partly overlapping:

- **Passengers and cargo forwarders**: Some passengers and cargo forwarders are part of the regional economy and local community, but some passengers and cargo forwarders are in the area in transit and have a very limited direct link to the regional economy.
- **Noise-affected homeowners**: The noise-affected homeowners are only a subset of all residents in the catchment area.

- **Airport suppliers**: The airport suppliers are a subset of the region's businesses.

Lastly, the diagram illustrates the multiple sources of airport capital funding.

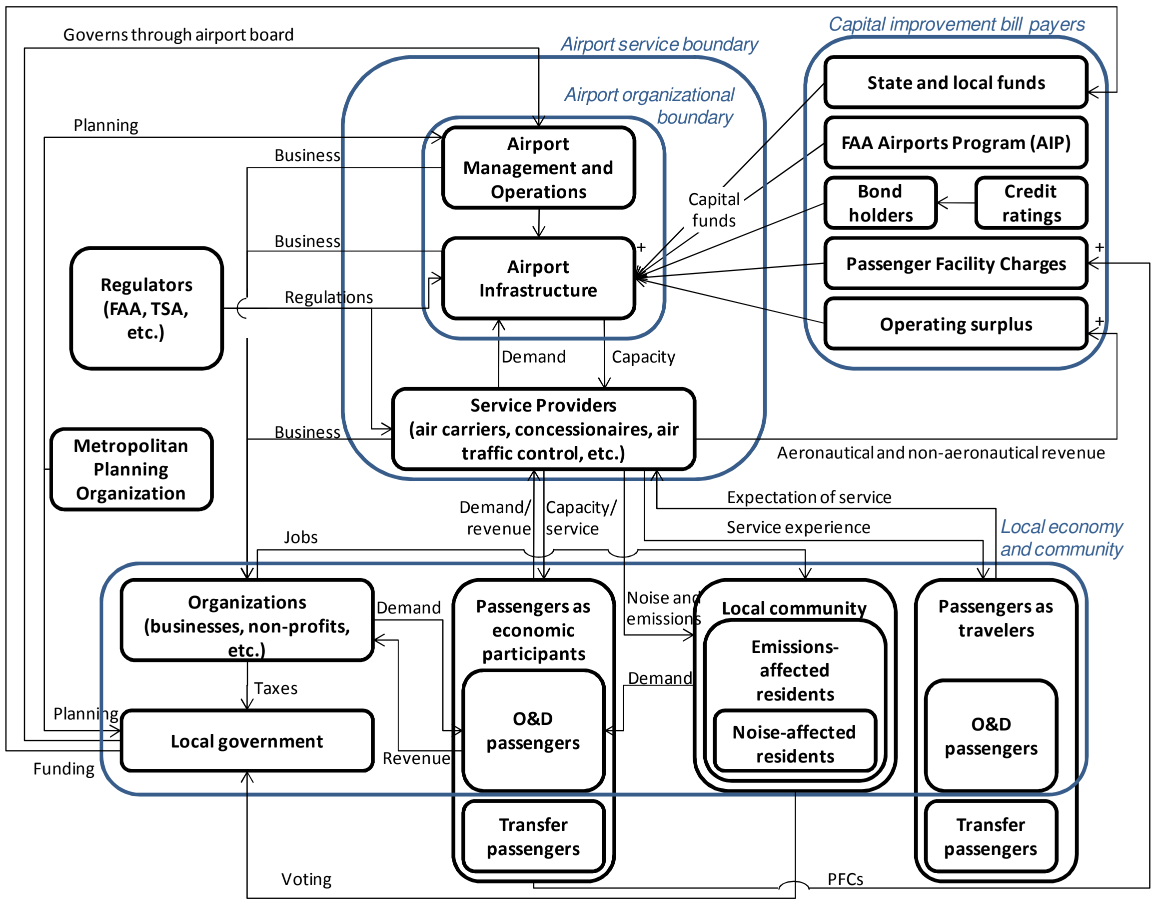

Figure 2.5 – Financial, Customer, and Other Relationships Between Airport Stakeholders

2.1.3.4.2 Airport Boundaries

Two different boundaries around the airport are identified in the diagram: The airport organizational boundary and the airport service boundary.

The airport's organizational boundary shows the limits of what is controlled by airport management. The boundary shows that airport management, and by extension, the airport board, can only control matters that relate to the design and configuration of airport infrastructure and the operational procedures and efficiency of its own organization. By contrast, the airport has limited control over the services provided at the airport such as the volume and types of air service and the types and quality of airport concessions.

This limited control is of relevance when contrasted with the airport service boundary. The airport service boundary represents the service of the airport as a function irrespective of the organizational responsibility for provisioning that service. For stakeholders outside the airport organization, the airport's performance may be judged not only on parameters within management's control, but also by factors such as what aircraft delays are or the frequency of services at the airport.

The arcs crossing the airport service boundary can be considered inputs to and outputs from the airport system. One can consider the concept of attaching "sensors" at these intersection points to measure the broader performance of the airport service in terms of generation of jobs, output of pollution and noise, service to passengers, etc.

2.1.3.4.3 System Loops

Within the diagram, several loops can be identified. These are either positively reinforcing loops where increased activity in one node propagates to increased activity in other nodes, or negatively reinforcing loops where increased activity in one node propagates to limitations in activity in other nodes. Depending on the nature of the loop, the timing of the impact of the loop effect will vary. Table 2.3 presents the loops in the stakeholder model, along with a discussion of the characteristics of the loop and the timing of the loop effect. Following the table are a graphic depictions of each of the loops.

Table 2.3 - System loops in stakeholder model

Loop	Description	Timing of effect
1. Economic activity positively reinforcing loop Illustrated in Figure 2.6.	Increased passenger and cargo volumes results in economic growth. In return, greater economic growth results in increased passenger and cargo volumes.	The timing of this effect can be characterized as on-going as it is a continuous and reinforcing effect (Kindleberger 1997).

Loop	Description	Timing of effect
2. Airport infrastructure capacity and funding positively reinforcing loop Illustrated in Figure 2.7.	Increases in capacity at the airport results in increased activity, which creates increased revenues for the airport in various forms. That in turn provides funding for capacity increases, and those capacity increases permit further growth. This assumes the presence of demand that will use the increased capacity.	The planning horizon for airport infrastructure is lengthy and is dependent on projections about future growth in airport traffic. A typical time horizon would be 10 to 20 years (Neufville & Odoni 2003, p. 70).

Loop	Description	Timing of effect
3. Airport noise and emissions negatively reinforcing loop Illustrated in Figure 2.8.	From increases in airport traffic come increased noise and emissions from airport operations and service providers. That has a negative impact on the local community, which expresses its desires through voting in local elections. The elected local government appoints the board and sets the mission for the airport. Accordingly, noise and emissions may result in limitations on future growth in capacity and restrictions on operations at the airport and limited funding for airport expansion, thereby constraining the opportunity for further growth of the airport.	Operational restrictions on the airport arise as a result of decisions within the local jurisdiction precipitated by community reactions to airport activity and this has a time horizon of several years, in order to go through the process of a Notice of Proposed Rulemaking (Wolfe & NewMyer 1985, pp. 83-85). In contrast, the constraining impact on capacity increases at the airport shares the same time horizon as the positively reinforcing loop described in the previous bullet of 10 to 20 years (Neufville & Odoni 2003, p. 70).

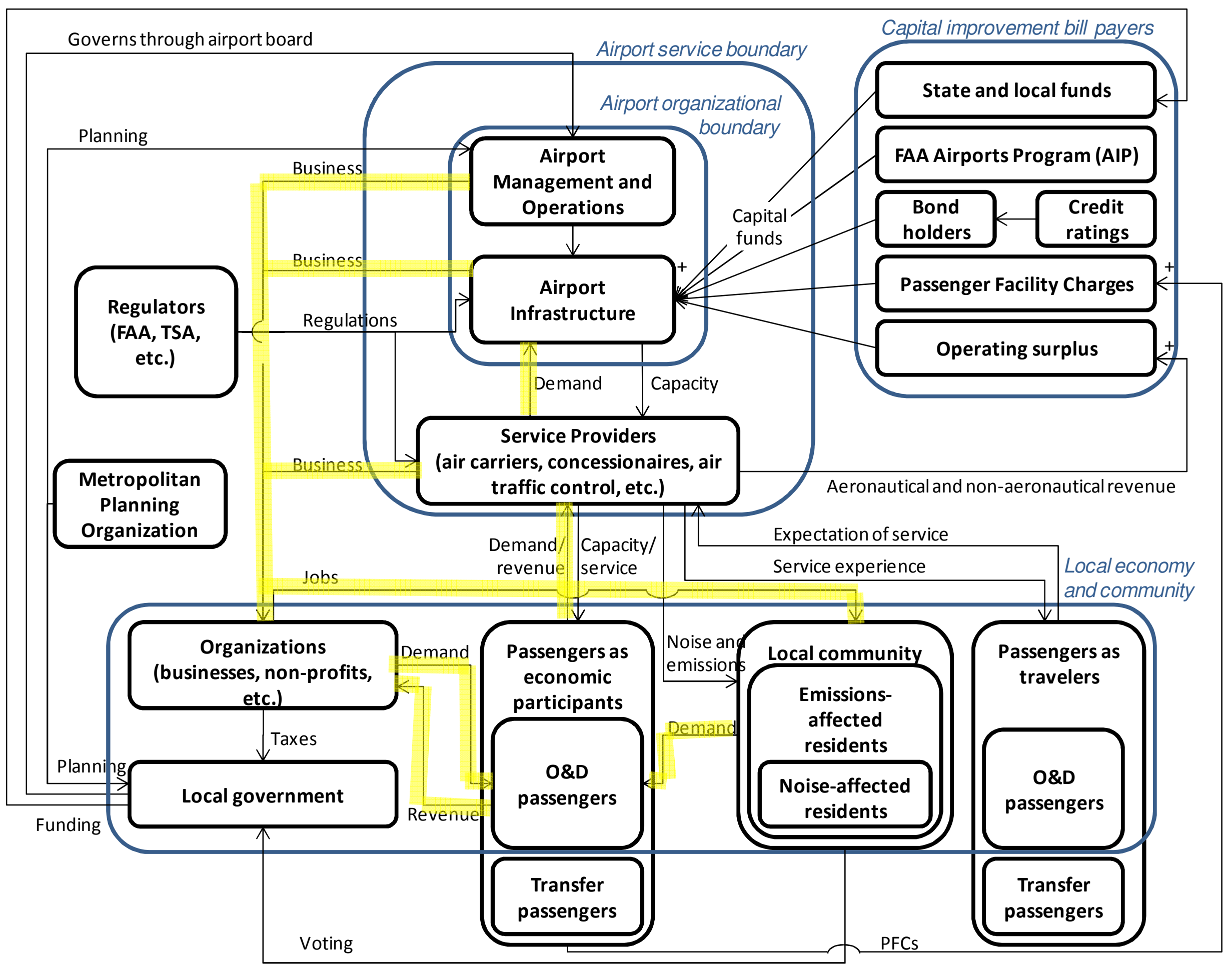

Figure 2.6 - Diagram of the economic activity positively reinforcing loop (highlighted in yellow)

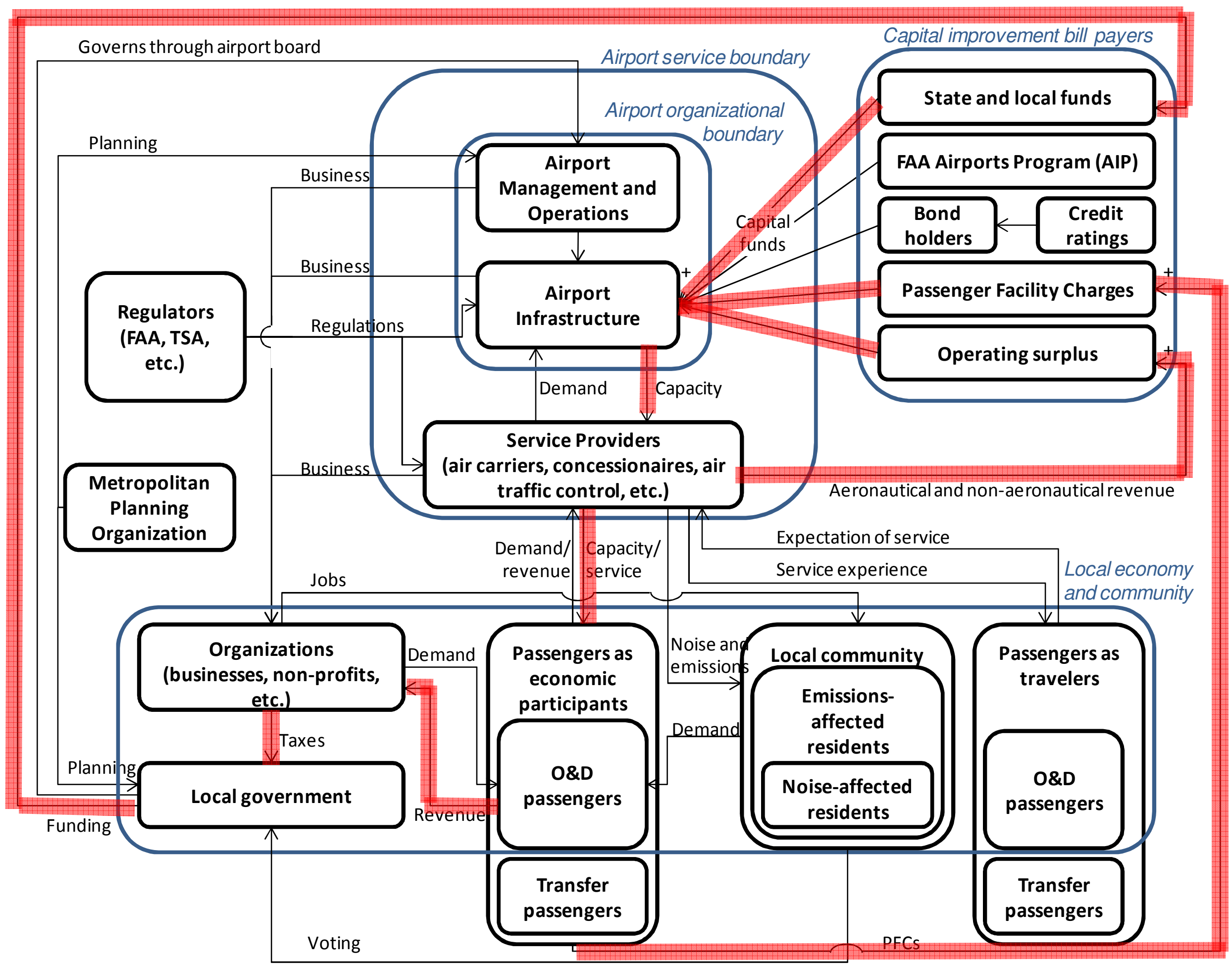

Figure 2.7 - Diagram of the capacity and funding positively reinforcing loop (highlighted in red)

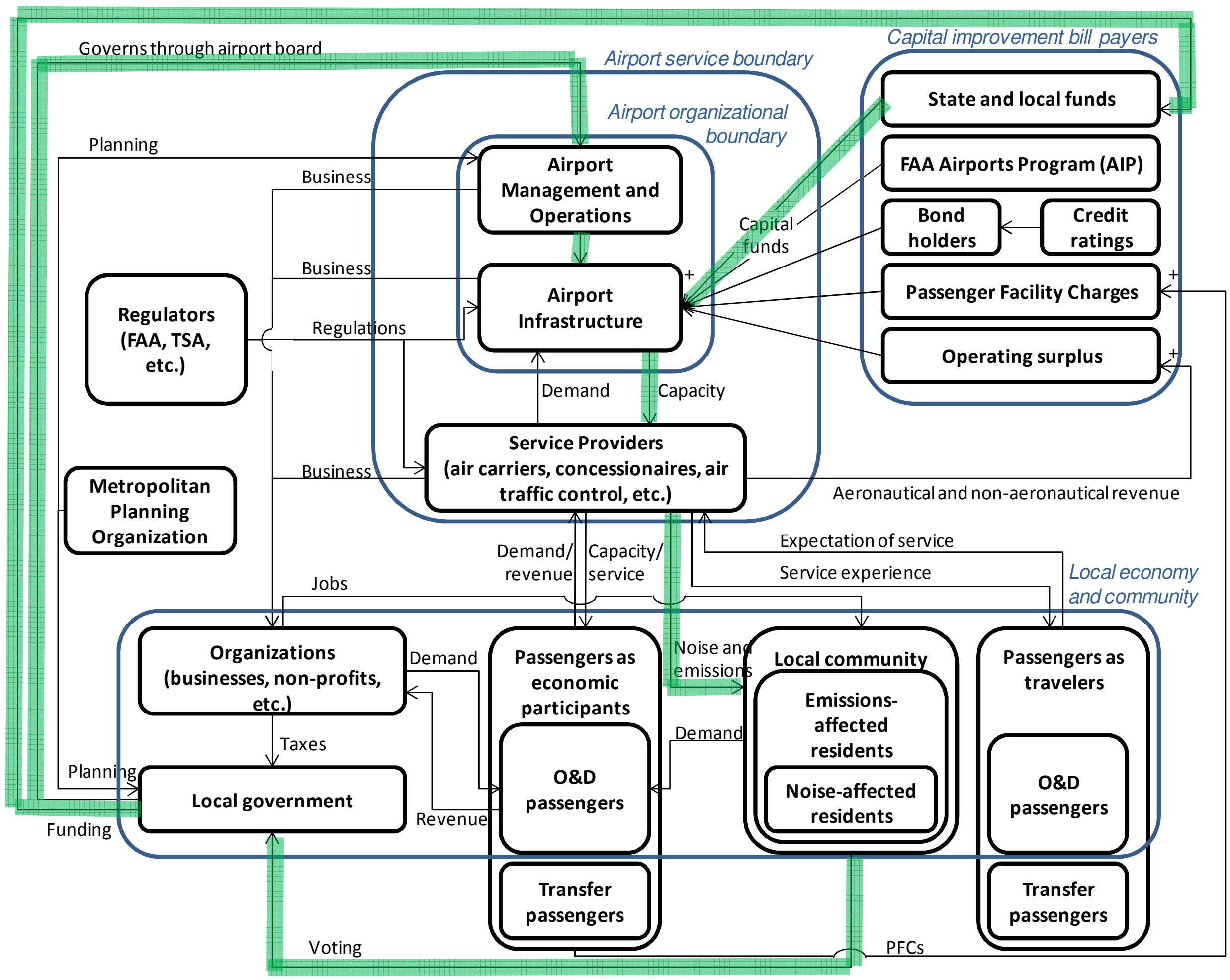

Figure 2.8 - Diagram of the noise and emissions negatively reinforcing loop (highlighted in green)

2.2 Airport Benchmarking

This section comprises an introduction to the concept of benchmarking, a discussion of analytical techniques for benchmarking, and a review of past academic and industry benchmarks. It finishes with a discussion of the choice of analytical models in airport benchmarks and a review of the implications of the new stakeholder model on past benchmarks.

2.2.1 Overview of Benchmarking

Benchmarking serves two purposes. The first is to measure the performance of one entity – organizational or otherwise – and compare it to the performance of other, similar entities. The second purpose is to identify those practices which enable the success of top performing entities; referred to as "best practices".

Robert C. Camp is widely credited as one of the originators of the practice of benchmarking in for example (Spendolini 1992) and (Yasin M. M. 2002). Camp pioneered the use of benchmarking at Xerox Corporation in the late 1970s and early 1980s as its photocopier business was being threatened by foreign competition (Camp 1989) (Camp 1993). Xerox ultimately implemented what is perhaps the first documented example of the full cycle of performance measurement, benchmarking, and best practice identification: Xerox measured the cost and quality of its products and compared these metrics to those of its competition. This allowed Xerox to identify the areas in which a performance gap existed, either in terms of cost or quality. To close this gap, Xerox studied the management practices of its competitors as well as organizations outside its industry, such as L.L. Bean, and was able to identify several key practices which drove those organizations' success.

Benchmarking has since been applied both in industry and academia in numerous studies. Dattakumar and Jagadeesh (Dattakumar & Jagadeesh 2003) conducted an extensive literature review and found more than 350 publications on the topic as of June 2002. By way of examples, Beretta et al. (Beretta et al. 1998) conducted benchmarking of accounting processes; Mann et al. (Mann et al. 1999) benchmarked the food and drinks industry; Matzko and Wingfield (Matzko & Wingfield 1995) described benchmarking of banks; Min and Min (Hokey Min & Hyesung Min 1996) benchmarked hotels; Ulusoy and Ikiz (Ulusoy et al. 2001) benchmarked manufacturing, and Zairi (Zairi 1998) benchmarked logistics.

Yasin (Yasin M. M. 2002) makes a distinction between competitive analysis and benchmarking by asserting that "although competitive analysis is useful in assessing one's position relative to the competition, it usually does not provide insights as to how competitors achieved this position, i.e. through what methods or processes. In contrast, the benchmarking process goes beyond comparison of results to include analysis of organizational processes and methods."

McNair and Leibfried (McNair & Leibfried 1992) provide a generic benchmarking process framework which breaks the process into the following steps:

1. Identify core issues: Determine the overall goals against which performance should be measured and determine the associated metrics. Determine potential drivers of these performance metrics.
2. Internal baseline data collection: Collection of performance metrics as well as data on current processes and practices
3. External data collection: Collection of the comparative dataset
4. Analysis: Identifying performance gaps and best practices which may be used to address those performance gaps
5. Change/Implement: Implementation of the benchmark recommendations

This framework highlights the process for an organization of identifying its goals, measuring its gaps against those goals, and finally identifying and implementing practices to address those gaps.

One of the primary criticisms of benchmarking is that by comparing an organization's practice to what others are already doing it becomes an exercise in catch-up; benchmarking can only ever help close the gap between laggards and top performers but will not enable an organization to leapfrog into first place (Thompson & Cox 1997). In response to that criticism, the argument has been made that innovation can in fact take place by looking at processes outside an organization's industry peers in areas where similar work is being performed; so-called functional benchmarking (Ahmed & Rafiq 1998).

2.2.2 Analytical Techniques for Benchmarking

This section provides an overview of the different analytical techniques available for benchmarking. It also presents an in-depth view of Data Envelopment Analysis (DEA) and its variations.

2.2.2.1 Overview of Methods for Calculation of Benchmark Scores

The benchmark score calculation is the process of analyzing the set of benchmark parameters to compute an overall assessment of performance for each participating enterprise, and ranking those enterprises from best to worst.

The factors that impact the choice of methodology for a benchmarking exercise are the number of benchmark parameters, the number of dimensions of the parameters, and whether or not a utility function for inputs and outputs is known.

Figure 2.9 shows a benchmarking model decision tree which summarizes the types of benchmarking models which will be reviewed in the subsequent sections.

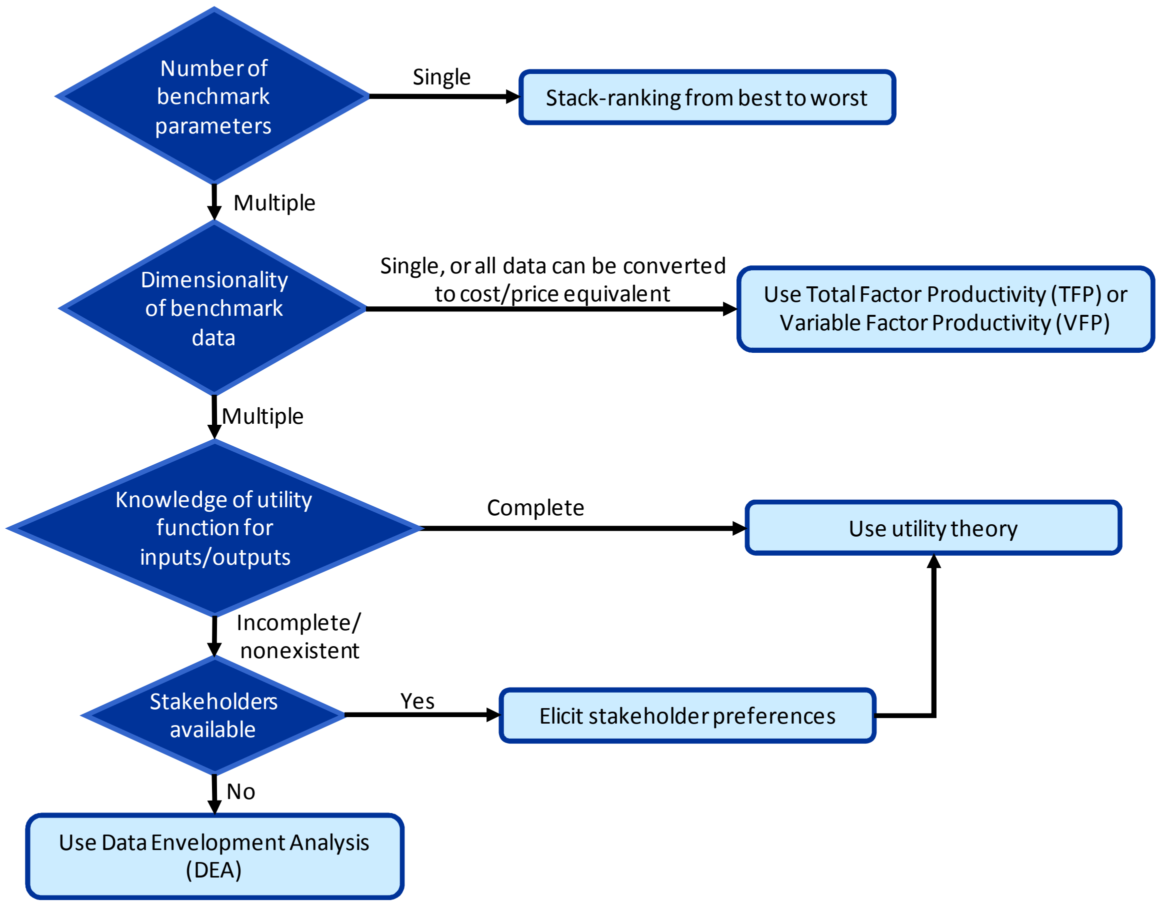

Figure 2.9 - Benchmarking Methodology Decision Tree

2.2.2.2 Benchmarking with a Single Performance Parameter

An example of a benchmark study where the analyst is comparing the performance of different entities on a single parameter would be the percentage of flights that arrive on time. The analyst may stack-rank each participating enterprise from best to worst or the analyst may compute percentile-based assessments for each enterprise.

Benchmark studies may use measurements that are based on a ratio of some form of inputs or resources to the production of outputs. For instance in an auto manufacturing situation this ratio might be the assembly labor cost per car. Although this situation uses two parameters (the labor cost and the number of cars), the computation of a ratio of the two measures means that the benchmarking technique used is the same as for a single metric.

Computation of individual components of productivity performance has been referred to as single factor productivity (Tretheway et al. 1997, p. 99).

2.2.2.3 Benchmarking with Multiple Parameters of a Single Dimension

When the study involves more than one input or output parameter, a technique for combining several parameters into a single, overall assessment of performance is necessary.

With input parameters x_1, x_2, ..., x_n and output parameters y_1, y_2, ..., y_m, the analyst is looking for a ratio between these sets of inputs and outputs, and wants to rank the participating enterprises from best to worst.

When the inputs used are the form of resource quantities that can be translated into an equivalent cost, the Total Factor Productivity (TFP) can be computed. The TFP compares the total outputs with the total inputs (Tretheway et al. 1997, p. 94).

In the example given in the previous section, the assembly labor cost per car represents a partial productivity measure since it only captures one part of the assembly cost. The total cost of assembly also involves for instance assembly equipment costs and costs for the assembly facility. An example of a TFP measure would instead be the calculation of the total production cost per car.

The Variable Factor Productivity (VFP) method is an alternative to TFP which focuses only on the subset of factors that are variable in nature (Tretheway et al. 1997, p. 99). For example, although some component of depreciation costs for capital assets such as the car assembly facility would be included in a TFP, a VFP analysis would exclude those costs.

2.2.2.4 Benchmarking with Multiple Parameters of Multiple Dimensions with Known Utility Functions

The aim of utility functions for comparing alternatives (e.g. airports) has been described as follows: "Assign numbers to the various alternatives so that alternative x is preferred to alternative y precisely when x gets a higher number than y." Furthermore, "[t]he number assigned to an alternative is usually called its utility, and sometimes its worth, and the assignment is a utility function" (Roberts 1972, p. 126).

The approach to benchmarking entities using utility functions is similar in nature to the VFP and TFP approaches just described, with the difference that rather than translating values of inputs and outputs into monetary equivalents, utility is computed for inputs and outputs. However, in order for utility theory to be applied, complete knowledge of the utility functions for all parameters being considered is necessary, as well as the utility functions for combinations of multiple different parameters.

2.2.2.5 Elicitation of Stakeholder Utility Functions for Benchmarking

When the benchmark includes multiple parameters in which utility functions are not known a priori but access to stakeholder representatives is available, it is possible to elicit the stakeholders' preferences for the mix of inputs and outputs.

Developed as part of work on recommender systems, the approach in (Alodhaibi et al. 2010) to elicitation of stakeholder preferences is applicable to benchmarking when stakeholder preferences are unknown. In the approach, a utility vector $\vec{u} = (u_1, u_2, .., u_n)$ is described, in which u_i is the utility of the value for each attribute under consideration (the attributes could be for example on-time performance and cost of operations at an airport). The utility for each attribute value would be extracted through the stakeholders' domain knowledge.

Second, each stakeholder's preferences for the mix of attributes is described in a vector of weights $\vec{w} = (w_1, w_2, .., w_n)$, where $\left|\vec{w}\right| = \sqrt{\sum_{i=1}^{n} {w_i}^2 = 1}$. The utility is then computed as $U_{\vec{w}}(\vec{u}) = w_1 u_1 + w_2 u_2 + .. + w_n u_n = \vec{w} \cdot \vec{u}$. The utility can be computed individually for a stakeholder or for a group of stakeholders, using composite values of $\vec{w}$.

To use this approach, it is necessary to have access to stakeholder representatives to elicit 1) their joint description of the utility of the range of possible values for each attribute and 2) their individual weights assigned to each attribute. With this knowledge, the analyst can compute benchmark results for the population of stakeholders as a whole, as well as benchmark results based on the preferences of individual stakeholders or sub-groups of stakeholders.

2.2.2.6 Benchmarking with Data Envelopment Analysis

Two challenges exist for benchmarking of enterprises across multiple parameters:

- **Parameter weighting**: How should the different parameters be weighted against one another? Continuing with the previously used example, it is possible that labor costs are more/less desirable than outsourcing costs, so is it an accurate representation to simply sum them and treat them equally? TFP and VFP do include parameter weightings, but the analyst must come up with a rationale for determining those weightings, and this will introduce subjective aspects to the benchmark results.

- **Multidimensionality and lack of utility function**: Instances exist where the set of inputs and/or outputs cannot simply be summed into an overall score due to different dimensions and the lack of knowledge about the utility function. For instance, some inputs may be in the form of values that can be expressed as costs, such as labor resources, while others cannot, such as the number of runways or landing capacity.

To address these issues, the non-parametric technique Data Envelopment Analysis (DEA) was been introduced. The premise of DEA is that each entity in the comparison set is a Decision-Making Unit (DMU) which has made conscious decisions about how to weigh the importance of each of the inputs and outputs. DEA is a linear programming-based technique which computes the optimal parameter weightings for each DMU under the constraint that those parameter weights cannot lead to efficiency scores greater than 1.0 for any of the participating DMUs in the study. This efficiency score LP is solved once for each of the participating DMUs, generating a set of optimal weightings for each DMU.

DEA was introduced by Cooper, Charnes, and Rhodes in 1978 (A. Charnes et al. 1978), giving the basic DEA algorithm its designation, CCR.

The objective is to identify the DMU(s) with the best inherent efficiency in converting inputs $x_1, x_2, \ldots, x_n$ into outputs $y_1, y_2, \ldots, y_m$. All other DMUs are then ranked relative to the most efficient DMU(s).

Model for DMU a:

$$\max h_a = \frac{\sum_r u_r y_{ra}}{\sum_i v_i x_{ia}}$$

Where u_r and v_i are weights applied to outputs y_{rj} and inputs x_{ij}

Subject to $$\frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1 \qquad \text{for each unit } j$$

$$u_r, v_i \geq 0$$

This fractional program is solved once for each of the DMUs, resulting in a set of optimal weights $u_1, u_2, \ldots, u_m$ and $v_1, v_2, \ldots, v_n$ for each DMU. However, the fractional program is difficult to solve, so the CCR creators proposed the following approach for turning it into a linear program:

Set the denominator in the objective function equal to some positive constant *c*. Move the equality of the denominator and *c* to the constraint section. Also cross-multiply the original constraint and rearrange it.

$$\max h_a = \frac{\sum_r u_r y_{ra}}{c}$$

$$\text{Subject to} \quad \sum_i v_i x_{ia} = c$$

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \le 0 \qquad \text{for each unit j}$$

$$u_r, v_i \ge 0$$

This is valid for all positive values of *c*. In the literature, the value for *c* is commonly set equal to 1, since that results in an easier-to-read problem. This results in the following modified problem.

$$\max h_a = \sum_r u_r y_{ra}$$

$$\text{Subject to} \quad \sum_i v_i x_{i_a} = 1$$

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \le 0 \quad \text{for each unit j}$$

$$u_r, v_i \ge 0$$

Next, the program is converted to its dual. This is advantageous from a solution efficiency perspective since the primary problem generally has a small number of variables but a large number of constraints. Turning the program into its dual generates a larger number of variables but fewer constraints.

Conversion of primal CCR problem to its dual:

Primal problem:

Max C [v, u] where C = [0_1,…,0_n, y_a]

Subject to A [v, u] ≤ B where $A = \begin{bmatrix} x_a & 0 \\ -X & Y \end{bmatrix}$

and $B = \begin{bmatrix} 1 \\ 0_1 \\ \cdots \\ 0_n \end{bmatrix}$

u, v ≥ 0

Dual problem:

Min [θ_a, λ] B where $\lambda = \lambda_1 \ldots \lambda_n$

Subject to [θ_a, λ] A ≥ C

θ_a unbounded, $\lambda \geq 0$

The expanded form of the dual CCR problem is:

$\min(\theta_a,\lambda) = \theta_a$ where λ is a vector $\lambda_1 \ldots \lambda_n$ and θ_a is a scalar.

Subject to
$$\theta_a x_a - X\lambda \geq 0$$
$$Y\lambda \geq y_a$$
$$\lambda \geq 0$$

Finally, slacks are added:

$$\min(\theta_a,\lambda) = \theta_a$$

Subject to
$$\theta_a x_a - X\lambda = s^-$$
$$Y\lambda = y_a + s^+$$
$$\lambda \geq 0, s^+ \geq 0, s^- \geq 0$$

The nonzero elements of the vector λ identify the active peers for each DMU; peers being those fully efficient DMUs that make up the reference set for any inefficient DMUs. Hence, in many examples with DMUs *A*, *B*, *C*, etc., the elements of λ are not identified as λ_1, λ_2, λ_3,… but instead as λ_A, λ_B, λ_C,… to indicate correspondence with each DMU.

For fully efficient DMUs, its corresponding element of λ will be equal to 1 and all other elements are equal to 0. For inefficient DMUs, one or more elements of λ that don't correspond to that DMU will be greater than 0.

A number of different variations of the basic CCR methodology have been proposed to address a number of perceived shortcomings. What follows is an overview of some DEA variations. These DEA variations are covered in (William Wager Cooper et al. 2006), (Abraham Charnes 1994), and (Ray 2004).

2.2.2.6.1 Adding a Positivity Constraint

The first variation to the basic CCR methodology was introduced by the original CCR authors (A. Charnes et al. 1979). This was a modification after it became clear that in the original CCR model, there were cases where some DMUs achieved full efficiency by ignoring some inputs/outputs by setting the corresponding weights to 0. The authors altered the original non-negativity constraint on the parameter weights to $u_r, v_i > 0$. This positivity constraint is generally implemented by the constraint $u_r, v_i > \varepsilon$ where ε is an infinitesimal constant.

2.2.2.6.2 BCC

The Banker Charnes Cooper (BCC) model is an extension to the original CCR model which introduces the concept of variable returns to scale (VRS) (R. D. Banker et al. 1984).

The concept of VRS is described in (William Wager Cooper et al. 2006) by way of a one-input and one-output example. The BCC model's frontier is in the form of a convex hull of the fully efficient DMUs. As a result of the convex hull, the number of fully efficient DMUs in a BCC model will always be greater than or equal to the number of fully efficient DMUs in a CCR model for the same data. An example from (William Wager Cooper et al. 2006) is recreated in Figure 2.10.

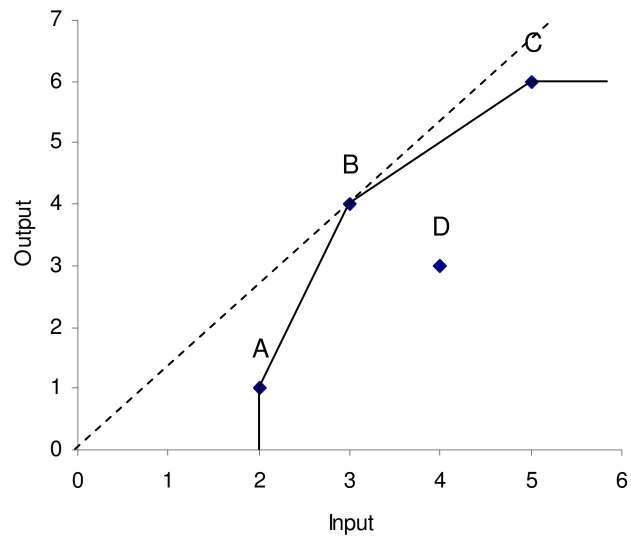

Figure 2.10 – BCC example

In this example, the dashed line represents the efficient frontier for the CCR model while the solid lines make up the BCC frontier. In the BCC model, A, B, and C are all fully efficient while only B is efficient in the CCR model. D is inefficient in both models. The slope of the BCC frontier lines shows that the model identifies increasing returns to scale between A and B since the BCC frontier on that segment has a steeper slope than the CCR frontier. Conversely, decreasing returns to scale exist on the segment between B and C for the BCC model. In contrast to the BCC model's VRS, the CCR model has an assumption of Constant Returns to Scale (CRS).

The BCC model is created by adding the below constraint to the dual problem.

$$\sum_{j=1}^{n} \lambda_j = 1$$

Because of the added constraint, the feasible region for the BCC model is always also part of the feasible region for the CCR model.

This constraint is also expressed as $e\lambda = 1$ where e is a unity vector $e_1 \ldots e_n$ with all elements equal to 1. The dual problem is expressed as follows:

$$\min(\theta_a, \lambda) = \theta_a$$

Subject to
$$\begin{aligned} &\theta_a x_a - X\lambda = s^- \\ &Y\lambda = y_a + s^+ \\ &e\lambda = 1 \\ &\lambda \geq 0, s^+ \geq 0, s^- \geq 0 \end{aligned}$$

In this model, input-oriented and output-oriented versions of the model must be distinguished. Input-oriented refers to the fact that the objective function is specified to minimize inputs while keeping outputs constant, while output-oriented refers to maximizing outputs while keeping inputs constant. In the CCR model, the solutions to input and output-oriented models are the same thanks to the constant returns to scale.

What is presented above is the input-oriented version of the BCC model. The output-oriented version of the model is similar in nature but its formulation focuses on maximizing outputs through the following formulation:

$$\max(\phi_a, \lambda) = \phi_a$$

Subject to

$$\begin{aligned} &\phi_a y_a - Y\lambda + s^+ = 0 \\ &X\lambda + s^- = x_a \\ &e\lambda = 1 \\ &\lambda \geq 0, s^+ \geq 0, s^- \geq 0 \end{aligned}$$

While the efficiency scores for an input-oriented and an output-oriented CCR model will be the same, this is not the case for the BCC model due to the convexity of the feasible region.

2.2.2.6.3 Additive

The additive model uses the slack variables directly in its objective function and combines both the input and the output-oriented components of the BCC models in its constraints. This model is referred to as a non-radial measure of efficiency since the objective function does not formulate a radial ratio based on the model frontier and the origin. The additive model was developed by Charnes, Cooper, Golany, Seiford, and Stutz (Ray 2004, p. 133).

The formulation of the model is:

$$\text{Max } z = e\,s^- + e\,s^+$$

Subject to
$$X\lambda + s^- = x_a$$
$$Y\lambda - s^+ = y_a$$
$$e\lambda = 1$$
$$\lambda \geq 0,\, s^+ \geq 0,\, s^- \geq 0$$

where e is a vector whose every element is 1

The efficiency score in the additive model is not unit independent, and slack for variables that have high magnitudes will have a greater impact on the efficiency score than those variables that do not.

2.2.2.6.4 Slacks-Based Measure of Efficiency and the Russell Measure

The Slacks-Based Measure (SBM) of efficiency was introduced by Tone (Tone 2001) and can be considered an extension to the additive model in that it is unit-independent when inputs and outputs of different dimensions are mixed.

The formulation of the SBM model follows:

$$\text{Min } z_a = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^{-} / x_{ia}}{1 + \frac{1}{n}\sum_{r=1}^{n} s_r^{+} / y_{ra}}$$

Subject to
$$X\lambda + s^{-} = x_a$$
$$Y\lambda - s^{+} = y_a$$
$$\lambda \geq 0, s^{+} \geq 0, s^{-} \geq 0$$

The unit independence is introduced by the fact that each slack in the objective function is divided by the corresponding input or output value for the DMU whose score is being computed. The constraints of this model are the same as those of the additive model.

The SBM formulation is a fractional problem. This conversion into a linear problem has been proposed (William W. Cooper et al. 2006, pp. 97-98):

$$\text{Min } \tau = t - \frac{1}{m}\sum_{i=1}^{m} ts_i^{-} / x_{ia}$$

Subject to
$$t + \frac{1}{n}\sum_{r=1}^{n} ts_r^{+} / y_{ra} = 1$$
$$X\lambda + s^{-} = x_a$$
$$Y\lambda - s^{+} = y_a$$
$$\lambda \geq 0, s^{+} \geq 0, s^{-} \geq 0, t > 0$$

New variables are defined:

$$S^{-} = ts^{-}, S^{+} = ts^{+}, \Lambda = t\lambda$$

Using the new variables, this linear program is formulated:

$$\text{Min } \tau = t - \frac{1}{m}\sum_{i=1}^{m} S_i^{-} / x_{ia}$$

$$\begin{aligned} \text{Subject to} \quad & t + \frac{1}{n}\sum_{r=1}^{n} S_r^{+} / y_{ra} = 1 \\ & X\Lambda + S^{-} = tx_a \\ & Y\Lambda - S^{+} = ty_a \\ & \Lambda \geq 0, S^{+} \geq 0, S^{-} \geq 0, t > 0 \end{aligned}$$

The optimal solution using the original variables are then obtained through:

$$z_a^{*} = \tau^{*}, \lambda^{*} = \frac{\Lambda^{*}}{t^{*}}, s^{-*} = \frac{S^{-*}}{t^{*}}, s^{+*} = \frac{S^{+*}}{t^{*}}$$

SBM shares near complete commonality with the Russell Measure of Efficiency (Färe & Knox Lovell 1978), and for that reason, the Russell Measure is not described further.

2.2.2.6.5 Free Disposal Hull

The Free Disposal Hull (FDH) model was introduced in 1984 (Marchand et al. 1984) and rests on the assumption that only observed combinations of inputs and outputs or combinations that are "worse" than those are feasible, removing the convexity assumption from the DEA model. Continuing with the previously used example, the shaded area in Figure 2.11 provides a visualization of the feasible region in the FDH model. While in the BCC model, any point along the segment stretching between points A and B would be feasible and on the fully efficient frontier, the FDH model eliminates any such points, creating a step function.

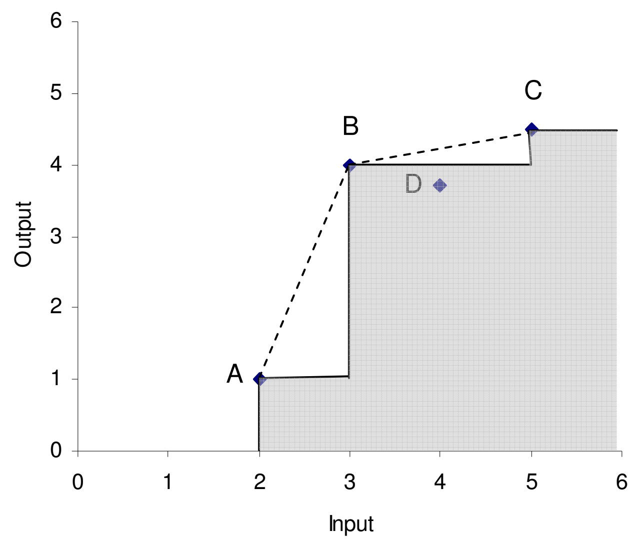

Figure 2.11 - Free Disposal Hull feasible region (shaded)

The feasible region for the FDH model is formulated as

$$P_{FDH} = \{(x, y) \mid x \geq x_j, y \leq y_j, x, y \geq 0, j = 1, 2, \ldots, n\}$$

The FDH model formulation in the input-oriented case is a mixed-integer program since the values in the λ vector are restricted to 0 and 1, and thanks to the unity constraint e λ = 1, only one of the values in λ is in fact nonzero.

$$\min(\theta_a, \lambda) = \theta_a$$

Subject to
$$\begin{aligned} &\theta_a x_a - X\lambda \geq 0 \\ &y_a - Y\lambda \leq 0 \\ &e\lambda = 1 \\ &\lambda_j \in \{0,1\} \end{aligned}$$

2.2.2.6.6 Super-efficiency

Several airport benchmarking studies, as will be shown in a subsequent section, use the concept of super-efficiency. However, two different meanings to the concept of super-efficiency were identified, even though both serve to break some of the ties that occur when too many DMUs are ranked as fully efficient.

In the super-efficiency version introduced by Andersen and Petersen (Andersen & Petersen 1993) and presented in (William Wager Cooper et al. 2006), super-efficiency refers to the removal of the values of the DMU for whom the efficiency score is being computed from the X and Y matrices in the constraints section. Those DMUs can in fact achieve efficiency scores greater than 1.0, and ties are (generally, but not always) broken.

This approach to super-efficiency can be applied to any number of DEA models, but is exemplified in the form of an input-oriented CCR model:

$$\min(\theta_a, \lambda) = \theta_a$$

Subject to
$$\theta_a x_a - \sum_{j=1, \neq a}^{n} \lambda_j x_j = s^-$$
$$\sum_{j=1, \neq a}^{n} \lambda_j y_j = y_a + s^+$$
$$\lambda \geq 0, s^+ \geq 0, s^- \geq 0$$

In contrast to this model, one airport benchmarking paper (Bazargan & Vasigh 2003) which will be reviewed in a later section refers to the use of a "super-efficient" DMU in a different manner. The paper uses a "super-efficient" DMU in a CCR study in order to break ties between too many fully efficient DMUs. In their approach, an "artificial" DMU is introduced to the study by assigning it the most favorable inputs and output values present in the study, drastically reducing the number of fully efficient DMUs. In most instances, the only fully efficient DMU in the study is the artificial, super-efficient DMU.

The method for selecting the inputs and outputs for the artificial DMU is described as:

$x_{iART} = \min(x_{ij}) \quad i = 1,....m, \; j = 1,....,n$

$y_{rART} = \min(y_{rj}) \quad r = 1,\ldots.s, \; j = 1,\ldots.,n$

2.2.2.6.7 Cross-Efficiency

DEA methods such as CCR and BCC have been called self appraisal as they allow each DMU to "choose" their own ideal set of input and output weights. In contrast to this self appraisal, Doyle and Green have proposed a form of peer appraisal they call cross efficiency (Doyle & Green 1994). Cross efficiency works by applying the ideal weights computed for all other DMUs to each DMU and computing an average efficiency score on that basis. The underlying assumption is that each DMU is evaluated based on the "opinions" of all of its peers about the relative importance of each parameter.

The base version of cross-efficiency (referred to as Simple Cross-Efficiency, of SXEF) can be determined by first computing the standard individual DEA scores using CCR, BCC, or some other DEA model. Doyle and Green introduce the idea of a rating DMU (when the DMU's own weights are used to assess the efficiency of other DMUs) and a rated DMU (when the weights of other DMUs are used to assess the efficiency of a DMU), and an example of their method is reproduced in Table 2.4.

Table 2.4 - Determination of SXEF (Doyle & Green 1994)

		Rated DMU						Averaged appraisal of peers
		1	2	3	4	5	6	
	1	E_{11}	E_{12}	E_{13}	E_{14}	E_{15}	E_{16}	A_1
	2	E_{21}	E_{22}	E_{23}	E_{24}	E_{25}	E_{26}	A_2
Rating	3	E_{31}	E_{32}	E_{33}	E_{34}	E_{35}	E_{36}	A_3
DMU	4	E_{41}	E_{42}	E_{43}	E_{44}	E_{45}	E_{46}	A_4
	5	E_{51}	E_{52}	E_{53}	E_{54}	E_{55}	E_{56}	A_5
	6	E_{61}	E_{62}	E_{63}	E_{64}	E_{65}	E_{66}	A_6
		e_1	e_2	e_3	e_4	e_5	e_6	

Averaged appraisal by peers

The scores e_1 through e_6 in Table 2.4 represent the SXEF scores.

Doyle and Green also introduce the concepts of Aggressive Cross-Efficiency (AXEF) and Benevolent Cross-Efficiency. These two models not only compute the most advantageous weights for each individual DMU but also seek to compute weights that respectively minimize and maximize the efficiency of all other DMUs.

The authors propose implementing this model in a two-phase method:

The first phase consists of computing the standard CCR or BCC efficiency score for each individual DMU. In the second phase, the objective function seeks to minimize (or maximize, respectively) the average efficiencies of all other DMUs, while imposing the constraint that the weights chosen for the DMU under consideration not worsen its own efficiency. This second phase is introduced based on the fact that the weights chosen in phase one may not be unique in achieving the same efficiency score for the DMU.

2.2.2.6.8 Malmquist Index

The Malmquist productivity index is used to measure productivity change over time. It includes two components: one measuring the individual DMU's productivity change between two time periods ("catch-up effect"); and one measuring the shift in the productivity frontier ("frontier-shift effect") (Chen & Iqbal Ali 2004).

The Malmquist index is not a DEA model in and of itself; rather, it incorporates the analyst's chosen DEA model in its computation of a DMU's performance change over time.

The Malmquist index is computed by multiplying the catch-up by the frontier-shift effect. To start, the catch-up effect is computed as follows, assuming we are comparing period 1 to period 2:

$$\text{Catch-up} = \frac{[\text{efficiency of } (x_a, y_a)_2 \text{ based on the period 2 frontier}]}{[\text{efficiency of } (x_a, y_a)_1 \text{ based on the period 1 frontier}]}$$

Each efficiency score is computed in separate DEA runs according to the selected DEA model. A catch-up greater than 1 indicates relative improvement between the two periods, while a value below 1 indicates a relative worsening of performance. A worsening of relative performance is possible in spite of absolute improvements of (x_a, y_a) since the peers in the benchmark may in fact have proportionately improved even more.

For the frontier-shift calculation, each of the observations of (x_a, y_a) are compared to the respective frontiers, as follows:

$$\phi_1 = \frac{[\text{efficiency of } (x_a, y_a)_1 \text{ based on the period 1 frontier}]}{[\text{efficiency of } (x_a, y_a)_1 \text{ based on the period 2 frontier}]}$$

$$\phi_2 = \frac{[\text{efficiency of } (x_a, y_a)_2 \text{ based on the period 1 frontier}]}{[\text{efficiency of } (x_a, y_a)_2 \text{ based on the period 2 frontier}]}$$

The frontier-shift is then computed as the geometric mean of the two ratios:

Frontier-shift = $\sqrt{\phi_1 \phi_2}$

The resulting Malmquist index is computed as (catch-up) * (frontier-shift)

Studies, e.g. (Sarkis 2000), have instead of a Malmquist index used repeated observations of the same DMUs over time as separate observations in a DEA analysis using the standard models such as CCR or BCC, thereby getting for instance 40 observations from 10 DMUs at four different points in time.

If a Malmquist index calculation is not done, but all observations are combined, the implicit assumption is that the conditions for achieving efficiency remain the same over time for individual DMUs and that no underlying factors (e.g. inflation, improved technology, etc.) have changed over time.

2.2.2.6.9 Radius of Classification Preservation

The measure of Radius of Classification Preservation (RCP) is a measure of the degree to which a DMU's inputs and outputs must be moved before the DMU's classification as efficient or inefficient changes (Rousseau & Semple 1995). A standard DEA model such as BCC or the additive model is used to compute initial efficiency scores. A maximum change ("radius") is then computed for each DMU to determine the sensitivity of its classification to changes in its inputs and outputs. This radius is proposed for use as a method for distinguishing between multiple efficient DMUs as determined by the initial DEA calculation.

For an efficient DMU, the RCP value is determined as follows:

$$\min \alpha^{+}$$

$$\text{subject to:} \quad \begin{aligned} & Y^{(E)}\lambda - s^{+} + \alpha^{+} e_s = y_E \\ & X^{(E)}\lambda + s^{-} - \alpha^{+} e_m = x_E \\ & e^T \lambda = 1 \\ & \lambda, s^{+}, s^{-}, \alpha^{+} \geq 0 \end{aligned}$$

where the subscript E indicates the efficient DMU whose RCP value is being calculated; $Y^{(E)}$ and $X^{(E)}$ are identical to the standard Y and X matrices with the exception that the columns corresponding to E are removed; and e, e_m, and e_s are vectors whose every element is 1.

2.2.2.6.10 Inefficiency Frontier

The concept of an inefficiency frontier is introduced in (Jiang et al. 2010). The authors observe that the general DEA models are based on some measure of the distance to the efficient frontier. They argue that of importance is not only a DMU's distance from the efficient frontier but also a DMU's distance from the inefficient frontier, with the inefficient frontier being defined as the convex hull of inefficient points. This concept of an inefficient frontier is illustrated in the one-input and one-output example in Figure 2.12. Points C, B, and D are on the inefficient frontier.

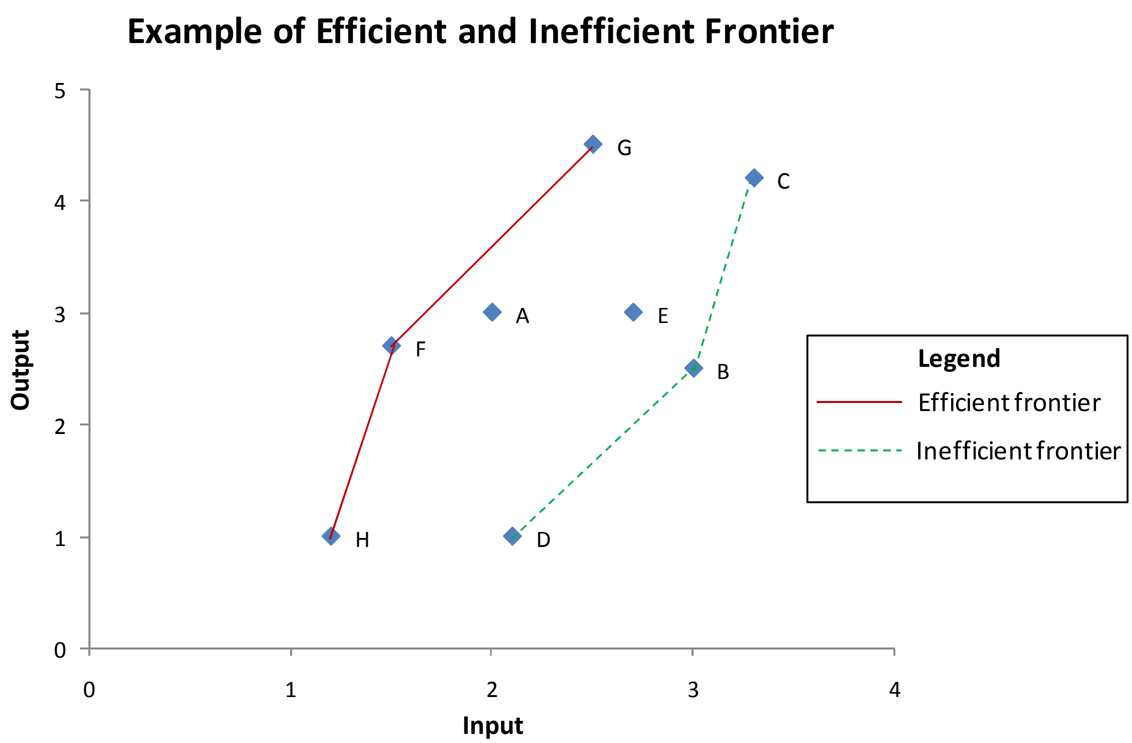

Figure 2.12 – Example of efficient and inefficient frontier

The formal definition of the strongly inefficient frontier includes the concept of an anti-production possibility set (APPS) which is the area outside of the production possibility set described in the constraints section of the general DEA models. The APPS is defined as follows:

$$APPS = \{(x, y) | x \leq X\lambda, y \geq Y\lambda, e^T\lambda = 1, \lambda \geq 0\}$$

The inefficient frontier (IF) is then defined as:

$$IF = \{(x, y) \in APPS | \forall (x', y') \in R_+^{m+s}, if (-x', y') \nleq (-x, y) \Rightarrow (x', y') \notin APPS\}$$

The authors then define a number of different means of computing the distance to the inefficient frontier that are similar to those described in previous sections for computing the distance to the efficient frontier. The distance to the inefficient frontier can then be used for breaking ties between DMUs.

2.2.2.6.11 Additive Model Adjusted for Negative Data

The models reviewed to this point are not able to incorporate input and output values that may take on negative values. For that reason, Pastor and Ruiz have proposed an adjustment to a model similar to the SBM/Russell measure of efficiency which results in a units-independent additive model which can treat negative inputs and outputs (Zhu & Cook 2007, p. 76).

The model specification follows:

$$\text{Max} \quad Z_a = \frac{1}{m+s}\left(\sum_{i=1}^{m}\frac{s_{i0}^{-}}{R_{i0}^{-}} + \sum_{r=1}^{s}\frac{s_{r0}^{+}}{R_{r0}^{+}}\right)$$

$$\text{Subject to} \quad X\lambda + s^{-} = x_a$$

$$Y\lambda - s^{+} = y_a$$

$$e\lambda = 1$$

$$\lambda \geq 0, s^{+} \geq 0, s^{-} \geq 0$$

$$\text{with} \quad R_{ra}^{+} = \max_{j=1,..,n}\{y_{rj}\} - y_{ra} \forall r, R_{ia}^{-} = x_{ia} - \min_{j=1,..,n}\{x_{ij}\} \forall i$$

where the slacks corresponding to any R_{i0}^{-} or R_{r0}^{+} with a value of zero are ignored.

2.2.2.6.12 Integer Constraints in DEA

The DEA models presented thus far assume that all inputs and output parameters take on continuous values. However, in some contexts, some or all inputs and outputs are indivisible. For instance, the number of runways, aircraft, or terminal buildings can only take on integer values.

Lozano and Villa (Zhu & Cook 2007, pp. 271-288) provide implementation guidance for applying integer constraints to several typical DEA models but also caution that this guidance does not apply to certain DEA variants (Zhu & Cook 2007, p 288).

The adjustments to the DEA models for incorporating integer constraints vary, but a typical example is the variation of the CCR input-oriented model with integer constraints:

$$\min(\theta_a,\lambda) = \theta_a$$

Subject to
$$\sum_j \lambda_j x_{ij} = x_i \forall i$$
$$\theta_a x_{ia} - s_i^- = x_i \forall i$$
$$\sum_j \lambda_j y_{kj} = y_k \forall k$$
$$y_{ka} + s_k^+ = y_k \forall k$$
$$\lambda \geq 0, s^+ \geq 0, s^- \geq 0, x_i \geq 0 \forall i, y_k \geq 0 \forall k$$
$$x_i \text{integer } \forall i \in I', y_k \text{integer } \forall k \in O'$$

where I' and O' are the input and output variables with integer constraints.

2.2.3 Review of Airport Benchmarks

This section presents a review of past benchmarks of airport performance. Two types of benchmarks can be identified: Those that were conducted for academic purposes and those conducted by industry organizations or consultancies. The former category of benchmarks is available through academic publications, giving insight into the metrics and model used. In contrast, the latter category is often available only to member organizations or those able to pay the cost for accessing the benchmark results, making evaluation of the metrics and model used impossible in many cases.

The following subsections review the two types of benchmarks.

2.2.3.1 Review of Academic Airport Benchmarks

Airport benchmarking studies have proliferated over the past decade and a half. Airport benchmarking studies began appearing in the mid-1990s, for example by Hooper and Henscher (Hooper & Hensher 1997) and Gillen and Lall (Gillen & Lall 1997). Since then a number of studies have appeared, using different methodologies and covering different geographic areas.

This section starts with Table 2.5 which contains a summary listing of airport benchmarking studies in the literature, along with their characteristics on the input and output metrics used and the geographic region covered.

Subsequent to the summary table, each paper is discussed in terms of its method of analysis, whether any best practices, controllable factors, or investment strategies were identified in the study through post-ranking analysis (post-ranking analysis being analysis whereby the benchmark "scores" are used dependent variables to understand what might be driving them), and any criticism of the approach in other papers.

Table 2.5 - Overview of past airport benchmarks

Study	Inputs	Outputs	Geography
Size Versus Efficiency: A Case Study of US Commercial Airports (Bazargan & Vasigh 2003)	Operational costs; non-operational expenses; number of runways; number of gates	Passenger throughput; aircraft movements; aeronautical revenue; non-aeronautical revenue; percentage of on-time operations	U.S.
Relative Efficiency of European Airports (Pels et al. 2001)	<u>Air transport movement study (part 1)</u>: Airport surface area; total length of runways; number of aircraft parking positions at terminals; number of remote aircraft parking positions	Aircraft movements	Europe
Relative Efficiency of European Airports (Pels et al. 2001)	<u>Passenger movement study (part 2)</u>: Terminal size; number of aircraft parking positions at terminals; number of check-in desks; number of baggage claim belts	Passenger throughput	Europe

Study	Inputs	Outputs	Geography
Developing Measures of Airport Productivity and Performance (Gillen & Lall 1997)	Terminal efficiency study (part 1): Number of runways; number of gates; terminal area; number of employees; number of baggage collection belts; number of public parking spots	Passenger throughput; cargo throughput	U.S.
Developing Measures of Airport Productivity and Performance (Gillen & Lall 1997)	Aircraft movement study (part 2): Airport surface area; number of runways; runway area; number of employees	Aircraft movements	U.S.
Performance Based Clustering for Benchmarking of US Airports (Sarkis & Talluri 2004)	Operational costs; number of employees; number of gates; number of runways	Operational revenue; passenger throughput; aircraft movements; cargo throughput	U.S.

Study	Inputs	Outputs	Geography
Measuring Airports' Operating Efficiency: A Summary of the 2003 ATRS Global Airport Benchmarking Report (Oum & Yu 2004)	Number of employees; number of runways; number of gates; terminal area; purchased goods, materials, and services (outsourcing)	Passenger throughput; cargo throughput; aircraft movements; non-aeronautical revenue	Global
An application of DEA to measure the efficiency of Spanish airports prior to privatization (Martín & Román 2001)	Labor expenditure; capital expenditure; materials expenditure	Aircraft movements; passenger throughput; cargo throughput	Spain
Measuring Total Factor Productivity of Airports - An Index Number Approach (Hooper & Hensher 1997)	Capital expenditure	Aeronautical revenue; non-aeronatical revenue	Australia

Study	Inputs	Outputs	Geography
Measuring Airport Quality from the Airlines' Viewpoint (Adler & Berechman 2001)	Airport charges; minimum connecting times; number of passenger terminals; number of runways; distance to nearest city center	Level of satisfaction from the airline users of each airport	Primarily Western Europe
An analysis of the operational efficiency of major airports in the United States (Sarkis 2000)	Operating costs; number of employees; number of gates; runways	Operating revenue; airline aircraft movements; general aviation aircraft movements; passenger throughput; freight throughput	U.S.
Managerial Efficiency of Brazilian Airports (Pacheco & Fernandes 2003)	Number of employees; payroll costs; operating expense	Passenger volume; cargo volume; operating revenue; commercial revenue; other revenue	Brazil
Assessing efficiency of European airports: a total factor productivity approach (Nyshadham & Rao 2000)	Capital cost; labor cost; other cost	Workload units (normalized sum of passenger and cargo volumes); aeronautical revenue; non-aeronautical revenue; employees; assets	Europe

Study	Inputs	Outputs	Geography
The performance of BAA before and after privatization (Parker 1999)	Number of employees; capital costs; operating costs	Revenue; passenger volume; cargo and mail volume	UK
Airports in Argentina: Technical efficiency in the context of an economic crisis (Barros 2008)	Number of employees; runway area; airport apron area; passenger terminal area	Aircraft movements; passenger volume; cargo volume	Argentina
Performance evaluation of Italian airports: A data envelopment analysis (Barros & Dieke 2007)	Labor costs; capital costs; operational (non-labor) costs	Aircraft movements; passenger volume; cargo volume; aeronautical revenue; handling revenue; commercial revenue	Italy

2.2.3.1.1 Analysis of Individual Studies

This section analyses each of the studies listed in the table in the previous section. Each paper is reviewed for its method of analysis and whether post-ranking analysis uncovered any best practices, controllable factors, or investment strategies.

The literature was also studied to understand what criticism has been levied against this set of papers by other studies. With only one exception, none of the papers in the study could be found to have received any criticism from other authors.

Note that this section discusses the benchmark model used and makes frequent reference to DEA technique. DEA is discussed in detail in section 2.2.2.6.

Size Versus Efficiency: A Case Study of US Commercial Airports (Bazargan & Vasigh 2003)

This study uses a large number of inputs and outputs which describe the level of activity at the airport as well as its costs and revenues. No discussion is provided as to why these metrics were selected.

The study uses the basic DEA model, CCR, without any mention or consideration of alternate DEA models from the literature. Instead the analysis uses a new method of introducing a super-efficient artificial DMU with the lowest input levels and the highest output levels in order to “force-rank” all airports.

In terms of post-ranking analysis, the study uses the calculated DEA scores to compute the differences in performance between small, medium, and large hubs, and finds that small hubs are consistently more efficient than large hubs.

The only source of criticism of this paper is a study (Schaar & Sherry 2008) which found that if some of the underlying modelling choices had been different, the conclusion of the paper could have been completely reversed (large hubs being more efficient than small hubs).

<u>Relative Efficiency of European Airports</u> (Pels et al. 2001)

This study is composed of two sub-studies, one focused on airside performance, and one focused on terminal operations. Its primary method of analysis is the DEA BCC variety which takes into account economies of scale, and this methodology is selected after a relatively exhaustive analysis. Subsequently, the analysis also uses stochastic production frontier analysis to confirm the primary findings, which are that airports do operate under increasing returns to scale. The paper notes that much further work is needed to explain which factors drive more efficient performance, but does not venture into actually doing so.

<u>Developing Measures of Airport Productivity and Performance</u> (Gillen & Lall 1997)

As with the previous paper, this study includes one review of airside and one of terminal operations efficiency. The primary methodology is DEA, and this paper is one of the first studies to apply DEA to airport benchmarking. The study uses both CCR and BCC to compare and contrast the effects of assuming that economies of scale exist, and that they don't.

The authors undertake a thorough Tobit regression study for their post-ranking analysis, using the efficiency scores as the dependent variables, and a series of controllable and uncontrollable factors as independent variables to understand their impact. This is one of very few studies to go to these lengths in this type of analysis and the authors uncover several findings:

The authors identify that on the airside, having hub airlines and expanding gate capacity has significant impact on improving efficiency, and reducing general aviation's portion of operations also improves efficiency.

On the terminal side, the authors find that efficiency is improved by expanding the number of gates and managing them in such a way as to maximize their utilization.

<u>Performance Based Clustering for Benchmarking of US Airports</u> (Sarkis & Talluri 2004)

The authors of this study use both the basic DEA CCR model but also expand into the alternative aggressive cross-efficiency model (AXEF), which applies each set of parameter weightings across all of the airports in the study and then computes a mean value for each airport.

The authors don't venture very far into explanatory analysis of why some airports are more efficient than others, but observe that many high-performing airports fall into warm or stable weather areas. This finding is in line with (Sarkis 2000) which more systematically finds that so-called "snow-belt" airports will have the worst performance.

Measuring Airports' Operating Efficiency: A Summary of the 2003 ATRS Global Airport Benchmarking Report (Oum & Yu 2004)

This study is a summary of an industry benchmark from ATRS which is primarily concerned with airports' financial performance. The study uses Variable Factor Productivity (VFP) to compare airport performance. The authors provide limited details on how this methodology is used, but discuss the fact that the measures considered are focused directly or indirectly on financial performance.

The authors do make a point of normalizing efficiency scores (using the VFP method) for several factors that are considered outside the control of airport management, namely: airport size, average aircraft size, percentage of international traffic, percentage of air cargo in total traffic, and capacity constrained airports. The study also considers several factors that are within the control of management, such as portion of non-aviation revenue, level of outsourcing, and overall passenger satisfaction.

Among the uncontrollable factors, the authors find that airport size, percent cargo, and capacity constraints all have statistically significant positive coefficients, indicating that they all improve performance. They also find that the percent of international traffic has a statistically significant negative coefficient. Among the controllable factors, the authors find that the percentage of non-aviation related revenue and degrees of outsourcing both have statistically significant positive coefficients, indicating that they both help improve efficiency.

An application of DEA to measure the efficiency of Spanish airports prior to privatization (Martín & Román 2001)

The authors in this study compute both the CCR and BCC versions of DEA after a thorough discussion of the two. The authors find general evidence of the existence of economies of scale, but do not take their analysis any further as far as finding other factors that explain strong or poor performance.

Measuring Total Factor Productivity of Airports - An Index Number Approach (Hooper & Hensher 1997)

The authors of this study use index number Total Factor Productivity (TFP) which is a method for comparing an index of outputs to an index of inputs by computing the ratio between the two. The index number TFP method uses predetermined parameter weights. However, the index number TFP study present here only considers financial measures of performance; this means effectively studying the airport as an enterprise whose mission it is to maximize revenue and minimize costs.

The authors do not conduct any post-ranking analysis.

Measuring Airport Quality from the Airlines' Viewpoint (Adler & Berechman 2001)

This study – as the study title suggests – takes a different approach to measuring airport performance by examining airport quality from the point of view of airlines. The authors use Principal Component Analysis in order to reduce the number of output variables and thereby get better discrimination among the airports, and combine this technique with the introduction of a super-efficient airport.

However, the authors fail to discuss why the inputs to the model were selected and do not address the fact that the satisfaction scores do not scale with increased input volumes, which is a concern in a DEA analysis.

The authors find 1) that the amount of landing charges had little impact on quality performance; 2) that connection times have little impact on quality performance; and that 3) an increased focus on freight traffic does have some impact on quality scores.

An analysis of the operational efficiency of major airports in the United States (Sarkis 2000)

This study uses the largest variety of DEA methodologies to-date. It compares the results of each methodology but unfortunately does not discuss the appropriateness of each. The study includes the basic CCR and BCC models, but also four additional models:

- Simple cross-efficiency
- Aggressive cross-efficiency
- Ranked efficiency
- Radii of classification rankings

Using the findings from the full suite of different analytical methods, the study finds 1) that airports that are hubs for major carriers are more efficient than non-hubs; 2) that airports in multiple airport systems are not more efficient than single airport systems; and 3) that airports in snowbelts are less efficient than those not in snowbelts.

Managerial Efficiency of Brazilian Airports (Pacheco & Fernandes 2003)

This paper provides a short discussion on methodology, and lands on the BCC model with the motivation that it is dealing with airports of varying sizes. The study is limited in its analysis and does not enter into any discussion about drivers of performance. Instead, it compares its DEA scores with another, separate study, and creates a two-dimensional matrix of efficiency scores, but this does not result in further insights into the drivers of airport performance.

Assessing efficiency of European airports: a total factor productivity approach (Nyshadham & Rao 2000)

This study uses TFP and is focused on economic and productivity metrics, and takes a financial view of airport performance. It does not provide any insight into underlying factors that drive productivity.

The performance of BAA before and after privatization (Parker 1999)

This study uses time-series data to create a large number of observations in its DEA analysis. The author discusses both CCR and BCC and uses both in the analysis. The main purpose is to compare what happened to the British Airports Authority (BAA) performance after it was privatized.

The author is able to conclude that the impact of privatization was not measurable. As a side conclusion, he finds clear evidence of economies of scale.

Airports in Argentina: Technical efficiency in the context of an economic crisis (Barros 2008)

This study compares and contrasts the CCR and BCC methodologies, and ultimately lands on the fact that economies of scale are in fact present.

The study proceeds to conduct a regression analysis and similar to for example (Sarkis 2000), the author finds that hub airports are more efficient than non-hubs. The study also finds that in a time of economic crisis, smaller airports fared worse than larger airports.

<u>Performance evaluation of Italian airports: A data envelopment analysis</u> (Barros & Dieke 2007)

The authors start off using the CCR and BCC methodology but because of the low ratio of observations to variables, they find too many airports ranked as efficient. They proceed to using the cross-efficiency DEA as well as the super-efficient DEA models. However, the authors do not provide any deeper explanation for why one method is to be preferred over another.

The authors confirm that two different forms of economies of scale (airports with large capital assets, and airports with large workload volumes) exist. The authors also find evidence that privately managed airports perform better than those under partially private management.

2.2.3.1.2 Summary

The studies reviewed in the previous section generally share the common feature that they are largely focused on financially oriented metrics and productivity calculations. There appears to be general agreement on the presence of economies of scale in airport operations.

Nine of 14 studies include various forms of post-ranking analysis, generally through different types of regression analysis. These analyses have uncovered various controllable (e.g. weather) and controllable (e.g. outsourcing) factors that have an impact on performance.

In terms of models, DEA is applied in 11 of 14 studies, followed by some examples of Total Factor Productivity analysis. While all studies applying DEA use some form of CCR or BCC analysis, there is variation in the details of how these models are applied. Several studies use multiple versions of CCR and BCC analysis in the same study, and four studies also complement CCR and BCC with the use of other DEA models. There appears to be little consensus on which methods are most appropriate under which circumstances.

An area which in large part is absent in these studies relates to the selection of the inputs and outputs in the various studies: As evidenced in

Table 2.5, each of these studies, which have roughly the same objective - to measure the comparative efficiency of a group of airports - offers a different perspective on how to measure the efficiency of these airports. Part of the reason why these parameters were selected may be the availability of these values. In many countries, the United States included, public reporting requirements for airports makes for instance financial data and passenger and aircraft movement data relatively readily available. Although any analyst will be limited by the types of performance data available, many of the studies in the previous section do not discuss how this available data relates to the enterprise's performance goals.

2.2.3.2 Review of Industry and Association Airport Benchmarks

Several benchmarks are conducted by industry associations as well as for-profit corporations. These benchmarks are summarized in Table 2.6.

Table 2.6 - Industry Airport Benchmarks

Benchmark	Organization	Results publicly available
ACI-NA Airport Performance Benchmarking Program (Airports Council International - North America 2009b)	Airports Council International – North America (ACI-NA)	No
Rates and Charges Survey (American Association of Airport Executives 2006)	American Association of Airport Executives	Yes, but requires purchase
Global Airport Benchmarking Report (Air Transport Research Society 2009)	Air Transport Research Society (ATRS)	Yes, but requires purchase. Some analytical results published in academic journal as described in section 2.2.3.1.
Airport Performance Indicators (Jacobs Consultancy 2009b) Review of Airport Charges (Jacobs Consultancy 2009b)	Jacobs Consultancy	Yes, but requires purchase
North America Airport Satisfaction (J.D. Power 2008)	J.D. Power	Summary results available
Airport Service Benchmarking (Airports Council International 2009)	Airports Council International (ACI)	No

These benchmarks can be broadly classified into four categories:

- **Industry association benchmarks**: These benchmarks (Airports Council International - North America 2009b) (American Association of Airport Executives 2006) (Airports Council International 2009) are produced for airport management to "measure the performance of discrete airport functions" and to "increase efficiency, quality, and customer satisfaction." (Airports Council International - North America 2009b)
- **Consultancy benchmarks**: These benchmarks (Jacobs Consultancy 2009b) are produced by a firm that provides "planning and management consultancy services in transport and infrastructure, project procurement and investment appraisal" (Jacobs Consultancy 2009a).
- **Analyst benchmarks**: This benchmark (Air Transport Research Society 2009) is produced by an association of airport analysts. The purpose of this benchmark is to "is to measure and compare the performance of several important aspects of airport operation". No specific mention of the intended audience is made.
- **Consumer sentiment benchmarks**: This benchmark (J.D. Power 2008) surveys airport customer satisfaction. The purpose of this

> benchmark is to provide data that “companies worldwide use to improve their business” (J.D. Power 2009).

Although details about the metrics used and the benchmarking model employed are unavailable for these benchmarks, the general descriptions of the benchmarks provide some insight into their scope. Table 2.7 summarizes the categories of metrics used.

Table 2.7 - Summary of metrics used in industry benchmarks

Airport Benchmark	Aero-nautical charges/ revenues	Throughput (passengers, cargo, aircraft)	Non-aero-nautical revenues	Financial ratios	Quality of airport facilities and service	Operating and maintenance costs	Staff	Physical facilities (e.g. number of gates, runways, etc.)	Airside and landside processing efficiency	Quality of community airline service
	Category of Metrics									
ACI-NA Airport Performance Benchmarking Program (Airports Council International - North America 2009b)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Rates and Charges Survey (American Association of Airport Executives 2006)	✓	✓	✓					✓		
Global Airport Benchmarking Report (Air Transport Research Society 2009)	✓	✓	✓	✓		✓	✓			
Airport Performance Indicators (Jacobs Consultancy 2009b)	✓	✓	✓	✓		✓	✓			
Review of Airport Charges (Jacobs Consultancy 2009b)	✓									
North America Airport Satisfaction (J.D. Power 2008)					✓					
Airport Service Benchmarking (Airports Council International 2009)					✓					
Portion of benchmarks using this metric	71%	57%	57%	43%	43%	43%	43%	29%	14%	14%

The summary in Table 2.7 suggests that the industry benchmark produced by ACI-NA is the most comprehensive of the benchmarks in terms of its coverage. A common thread throughout most of the benchmarks is a focus on financial factors, although exceptions to this are the three benchmarks that focus on quality ratings as well as the benchmark that evaluates the quality of community airline service.

2.2.4 Implications of the Stakeholder Analysis on Airport Performance Benchmarking

Published airport benchmarks, as evidenced in section 2.2.2, do not account for the fact that U.S. airports function as public utilities and must address multiple stakeholder concerns. Benchmarking of U.S. airports should be grounded in the goals of their stakeholders. As the analysis in section 2.1.3.1 shows, the goals for the airport vary depending on the stakeholder, and the analysis shows that stakeholder goals sometimes conflict. For instance:

- Passengers want access to low fares while air carriers want access to high-yield markets.

- Residents in the local community want a minimum of noise and emissions but a number of other stakeholder groups want traffic to be maximized.

This analysis identifies two conceptual boundaries around the airport: 1) A boundary around the airport organization; and 2) a boundary around the airport service, which also includes service providers such as air carriers. The analysis shows that stakeholders who are located outside the airport service boundary have objectives whose fulfillment is not fully under the control of airport management.

Since U.S. airports function as public utilities, benchmarking of airport performance must be based on the goals of one or more airport stakeholders, and depending on the stakeholders included in the analysis conflicting goals may exist. Airport management must balance these sometimes opposing objectives for their stakeholders in determining performance goals.

The analysis also shows that not all aspects of stakeholders' performance goals for the airport are under the control of airport management. This is an important consideration in determining the right performance metrics for an airport performance benchmark and in interpreting the results of the airport benchmark.

A stakeholder-driven benchmark of airport performance can be a useful tool for determining in which airport improvement investments should be made since it can determine where the greatest benefits can be generated. The analysis shows that such a benchmark should be based on the goals of a number of airport stakeholders, and that it should not only be limited to factors within the direct control of airport management.

Similarly, benchmarks can be used to guide financial decisions about where to add or drop services for airport service providers. Such benchmarks must also be founded in the goals of those service providers when performance metrics are selected.

2.2.5 Analytical Techniques Used in Past Airport Benchmarks

Table 2.8 shows an overview of the benchmark models used in past benchmarks and illustrates that the primary model choice in those past studies has been DEA, but also shows that that a variety of different DEA models have been applied in past airport benchmark studies. This mix of model choices exist in spite of the fact that nearly all of these studies are some variation of a measure of how effectively these airports convert inputs such as labor and capital to desirable outputs such as passengers and aircraft movements. That a variety of different models have been applied to the same problem points to the need for a structured framework for benchmarking model selection.

Table 2.8 - Models used in past airport benchmark studies

Model:	VFP	TFP	DEA							
			CCR	CCR with super-efficiency	BCC	BCC with super-efficiency	AXEF	SXEF	RCP	BCC with Malmquist
Size Versus Efficiency: A Case Study of US Commercial Airports (Bazargan & Vasigh 2003)				X						
Relative Efficiency of European Airports (Pels et al. 2001)					X					
Developing Measures of Airport Productivity and Performance (Gillen & Lall 1997)			X		X					
Performance Based Clustering for Benchmarking of US Airports (Sarkis & Talluri 2004)			X				X			
Measuring Airports' Operating Efficiency: A Summary of the 2003 ATRS Global Airport Benchmarking Report (Oum & Yu 2004)	X									
An application of DEA to measure the efficiency of Spanish airports prior to privatization (Martín & Román 2001)			X		X					
Measuring Total Factor Productivity of Airports - An Index Number Approach (Hooper & Hensher 1997)		X								

Model:	VFP	TFP	DEA							
			CCR	CCR with super-efficiency	BCC	BCC with super-efficiency	AXEF	SXEF	RCP	BCC with Malmquist
Measuring Airport Quality from the Airlines' Viewpoint (Adler & Berechman 2001)						X				
An analysis of the operational efficiency of major airports in the United States (Sarkis 2000)			X	X	X		X	X	X	
Managerial Efficiency of Brazilian Airports (Pacheco & Fernandes 2003)					X					
Assessing efficiency of European airports: a total factor productivity approach (Nyshadham & Rao 2000)		X								
The performance of BAA before and after privatization (Parker 1999)			X		X					
Total factor productivity and efficiency of Australian airports (Abbott & Wu 2002)										X
Airports in Argentina: Technical efficiency in the context of an economic crisis (Barros 2008)					X					
Performance evaluation of Italian airports: A data envelopment analysis (Barros & Dieke 2007)						X	X			

2.3 Dissertation Problem Statement

The dissertation's problem statement is comprised of three components identified as gaps in this literature review. The following subsections summarize these three gaps.

2.3.1 Problem 1: Stakeholder and Goal Ambiguity

The review of past airport benchmarks in section 2.2.3 shows that past studies have not examined airport stakeholders and their goals in selecting the performance metrics in use in those benchmarks. Rather, it appears that many benchmarks are designed around data that was available to the researchers without any analysis of why and to whom this performance data was pertinent. As a result, the conclusions of those benchmarks lack relevance in relation to the true goals of the airport.

The analysis in section 2.1.3 shows that a range of different stakeholders with different and sometimes conflicting objectives exist. The complexity of the stakeholder model indicates that benchmarks must take a structured approach to selecting the stakeholders whose goals are to be reflected in the benchmark before making a determination about which metrics to include in the study.

2.3.2 Problem 2: Lack of Systematic Model Selection

The review in section 2.2.2 shows the existence of a variety of different DEA model variants and section 2.2.3.1 shows that past airport benchmarks have applied many of these DEA model variants. This variability in model choices exists in spite of the fact that all of the studies in section 2.2.3.1.1 review some version of the same problem: Which airports are most efficient? Meanwhile, an analysis has shown (Schaar & Sherry 2008) that the choice of benchmark model can have a radical impact on the results of the benchmark.

This suggests that to create valid benchmarks, a method is needed for systematically selecting a benchmark model, and that the choice about benchmark model must reflect the underlying characteristics of the domain being modeled.

2.3.3 Problem 3: No Benchmarks Apply a Systematic Process

As a result of the problems described in the previous two sections, no benchmark exists with a systematic approach to the selection of stakeholder goals and benchmark model. Analysis which will be described in section 4.1 shows that past benchmarks have consistently made poor selections in determining the DEA model in at least four categories, including the choice about which basic model for aggregating data into results to use, returns to scale, integer constraints, and the calculation of results across several time periods.

A set of new benchmark studies is necessary to address this gap in terms of systematically selecting performance metrics and analytical model in the benchmarks.

3 Chapter 3: Methodology

This methodology is designed to benchmark airports. The methodology addresses the problems identified in section 2.3: the lack of a systematic approach to selection of benchmark metrics based on stakeholder goals; the lack of a method for selecting the model for computing benchmark results; and the lack of benchmarks conducted using a systematic process.

This section describes a comprehensive step-by-step process for conducting a benchmark which not only computes benchmark results but also interprets those results and derives actionable findings.

Figure 3.1 provides an overview of the airport benchmarking methodology. Each step of the methodology is described in detail in the following subsections.

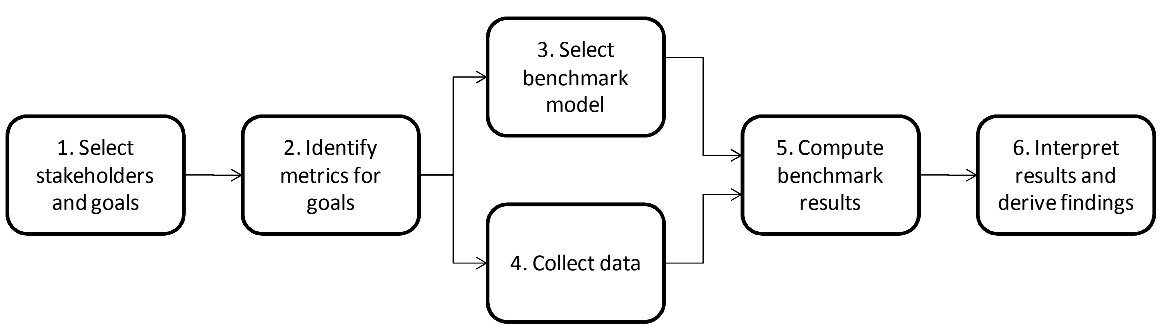

Figure 3.1 - Overview of airport benchmarking methodology

The steps in Figure 3.1 can be categorized into three phases, as described in Table 3.1.

Table 3.1 - Phases of benchmarking methodology

Phase	Description
Phase 1: Benchmark design	The design of the benchmark involves identifying the stakeholders and their goals, determining the metrics to be used, and selecting the DEA model for the benchmark. This is represented by steps 1, 2, and 3 in Figure 3.1.
Phase 2: Benchmark implementation	The benchmark implementation involves collecting the performance parameters and computing the benchmark scores using the DEA implementation, as shown in steps 4 and 5 in Figure 3.1
Phase 3: Analysis and interpretation of results	The analysis and interpretation of results has the objective of uncovering controllable and uncontrollable factors which impact the benchmark results.

The methodology was developed by creating a combination of leveraging and extending existing research with describing new methods of analysis. Table 3.2 presents an overview of the approach to developing each step in the methodology.

Table 3.2 - Development of benchmarking methodology

Step	Description of development approach
Step 1: Select stakeholders and goals	The model of stakeholders and their goals which provides the foundation for this step was based on an extensive literature survey and on knowledge elicitation sessions with 32 representatives of stakeholder groups.
Step 2: Identify metrics for goals	Identifying performance metrics which link to the stakeholder goals was found to be a process which had not previously been investigated by researchers, resulting in a completely new process being described in the dissertation.

Step	Description of development approach
Step 3: Select benchmark model	The framework for DEA model selection was based on the review of existing DEA models presented in section 2.2.2.6, and it extends the DEA framework described by (Kleine 2004). The development of heuristics for model selection using the DEA framework also leveraged the model review in section 2.2.2.6 as well as a number of sources relating to the operation of airports.
Step 4: Collect data	The data collection step was developed by conducting an inventory of data sources that pertain to the key aspects of airport operations and finance.
Step 5: Compute benchmark results	The software for computing DEA benchmark scores was developed using Matlab and C++, leveraging the CPLEX linear and mixed-integer program solver. The implementations were founded on the DEA model descriptions in section 2.2.2.6.

Step	Description of development approach
Step 6: Interpret results and derive findings	The structure of methods for interpreting results and deriving findings had not previously been addressed by researchers. The components of the structure were in part based on methods used in past benchmarks for analyzing results and on a review of statistical methods which do not require the normality assumption to hold.

3.1 Step 1: Select Stakeholders and Goals

As described in section 2.3.1, the benchmark design must have a foundation in stakeholder objectives for the airport. This step in the benchmark design phase takes one of two different starting points:

1. Starting with an analytical "angle" which describes the focus of the benchmark. For instance, such an angle could be "comparing the operational efficiency of airports" or "comparing the environmental or noise impact of airports". This approach may cover one or more stakeholder groups.

2. Starting with the definition of one or more stakeholder groups on behalf of which the benchmark will be conducted. For instance, this could be "a benchmark of airports' success in meeting the objectives of regional residents".

From this starting point, the stakeholder model and goals database described in section 2.1.3.3 is used to determine the goals on which the benchmark will be based.

If using the analytical "angle" described in item 1 above, the stakeholder model and goals database are used to identify the stakeholders that are relevant for this area and the subset of their goals that pertain to this domain are extracted. For instance, if the goal is to compare the environmental and noise impact of airports, the stakeholder model and database is used to identify all stakeholders that have goals which are relevant to environmental and noise performance, and these goals and their associated stakeholders are compiled to serve as the basis for the benchmark.

If using the stakeholder-focused definition of the benchmark, the full set of the stakeholders' goals for the airport are extracted from the database, and serve as the basis for the benchmark.

The resulting list of goals provides the starting point for the next phase of the benchmark methodology in which performance metrics are selected.

3.2 Step 2: Identify Metrics for Goals

Having selected the goals which should be reflected in the benchmark, select the metrics which reflect performance against these goals. There are two factors to be taken into consideration in assembling the performance metrics: i) Select the performance metrics and ii) analyze the metrics for any methodological discrepancies, and making any necessary adjustments to the metrics. This section describes these two sub-steps.

3.2.1 Selecting the Metrics

To select the metrics for a benchmark goal, complete the following two steps. The first step is to analyze the goal for whether it is at the "atomic" or composite level, by determining if it is composed of several elements which must be measured individually, or if it is specified in such a way that it can be described by a single metric. If the goal is composite, it is decomposed into its atomic elements.

Determining whether a goal is atomic or composite is based on a review of the literature in the area as well as knowledge elicitation sessions with stakeholder representatives. This body of research and expert knowledge gives the analyst insight about whether several components must be measured or if a metric which is all-encompassing is available. For instance, research about the environmental impact of airports may yield the insight that there are several component chemical compounds whose volume or concentration must be measured separately (FAA Office of Environment and Energy 2005). Conversely, research into the noise impact of traffic at an airport may yield the insight that the primary measure of concern is the number of residents within the area affected by more than a certain number of decibels of noise stemming from the airport traffic (Neufville & Odoni 2003, p. 178).

The second step is to select the individual metrics to use. In the ideal state any desirable performance data would be available, but in practice the benchmark is limited by the types of data available.

If the goal definition includes a specific metric (e.g. “maximize the total volume of passengers carried”), identify sources of data which most closely match that metric.

If the metric is more generally stated (e.g. “minimize delay”), the analyst starts by conducting an inventory of available relevant metrics (e.g. “aircraft arrival delay”, “aircraft departure delay”, “passenger arrival delay”, etc.). From this inventory, study the literature and interview stakeholder representatives to determine which one among the metrics most comprehensively addresses the goal.

The analysis of performance metrics may indicate that no available metric comprehensively addresses the goal. For instance, rather than data on total passenger volumes, data may only be available on the volume of domestic passengers. If this is the case, identify these limitations in the discussion of the benchmark results.

This metrics selection process is summarized in Figure 3.2.

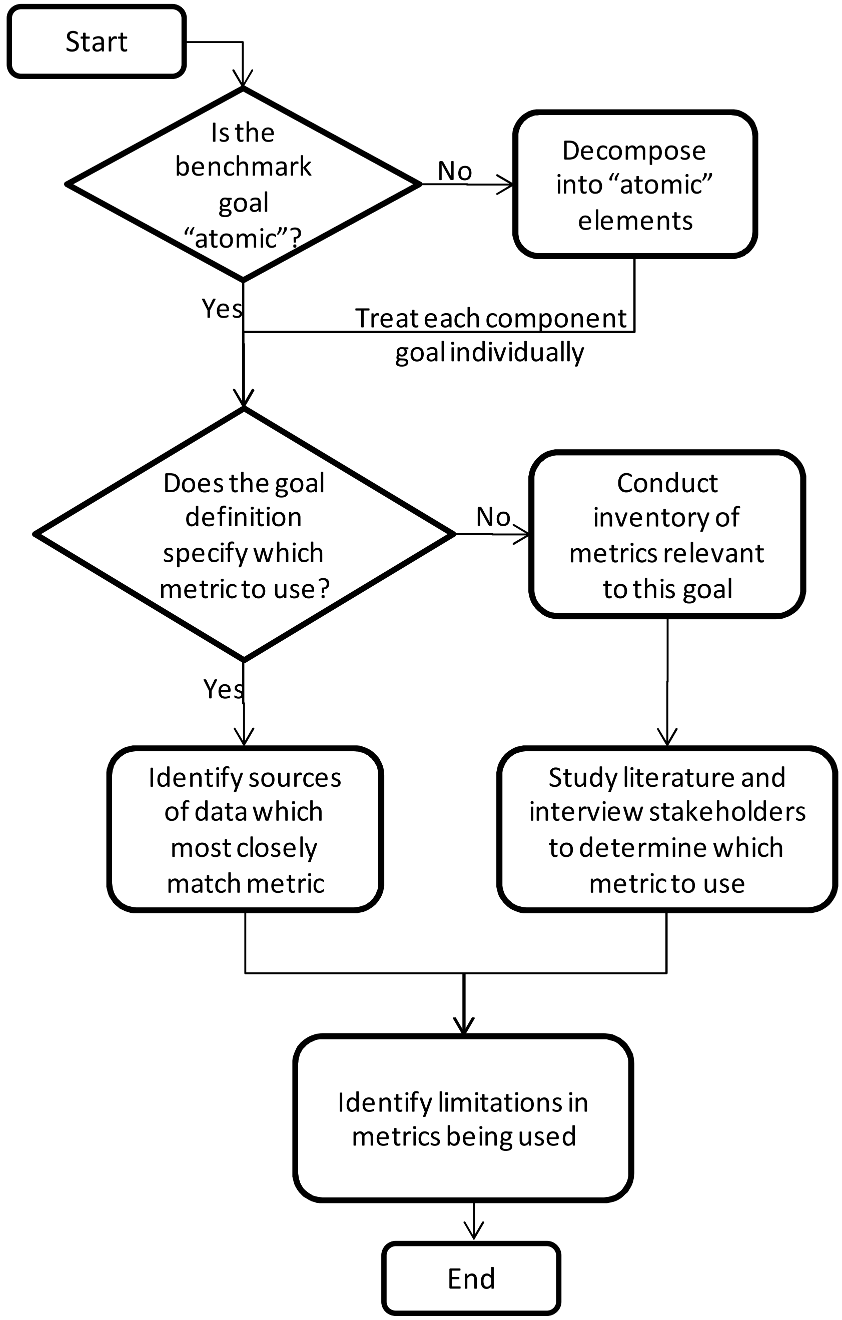

Figure 3.2 - Process for selecting metrics

3.2.2 Identifying Methodological Discrepancies

The step of identifying methodological discrepancies in the metrics that are selected serves to find cases where the metrics selected will not be suitable for use in a DEA analysis. The methodological discrepancies to avoid include:

- **Using too many metrics**: If too many metrics are used in the DEA run, the discriminatory power of DEA to separate the fully efficient DMUs from inefficient ones is reduced. (R. G. Dyson et al. 2001) suggests a rule of thumb that the number of DMUs should be greater than or equal to twice the product of the number of inputs and outputs, and proposes approaches to reducing the number of inputs and outputs if necessary. These approaches include converting inputs that can be priced into single cost values.
- **Combining indices and volume measures**: A DEA benchmark cannot mix indices such as percentages or other computed factors with metrics which scale with DMU size, such as passenger volumes. The DEA analysis must include only one type of measure.

3.3 Step 3: Select Benchmark Model

As discussed in section 2.2.2.6, several different DEA models exist and as shown in section 2.2.5, airport benchmark studies have applied many of these variations. This points to the need for a systematic approach to the selection of a DEA model. For airport benchmarking, this is accomplished by choosing the DEA model using the model selection framework along with the associated heuristics which are both presented in this section. The full details of the framework and heuristics as well as the background on their development are presented in Appendix A.

The DEA framework presents a structure for the choices that must be made in determining a DEA model for the benchmark. An overview of this framework is presented in the first subsection. The heuristics associated with the framework provide decision guidance for making selections in the framework when conducting an airport benchmark. A summary of the heuristics is presented in the second subsection. Combined, these two elements provide the modeler with a method for making a well-founded choice about which DEA model to use.

Figure 3.3 provides an overview of how the framework and heuristics should be applied.

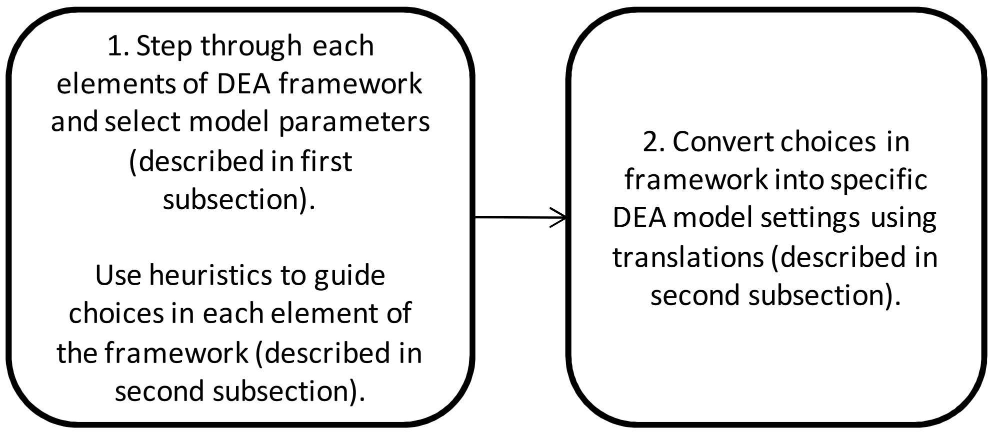

Figure 3.3 – Overview of application of DEA model framework and heuristics

3.3.1 A DEA Model Selection Framework

The framework for DEA model selection was extended from the work in (Kleine 2004). An overview of the framework is presented in Figure 3.4. In DEA modeling, step through each of the elements in the framework and make a choice for each element, based on the characteristics of the domain being modeled. For guidance on how to make selections in each element, use the heuristics presented in the next section. For full details on the definition of each element and value in the framework, refer to Appendix A.

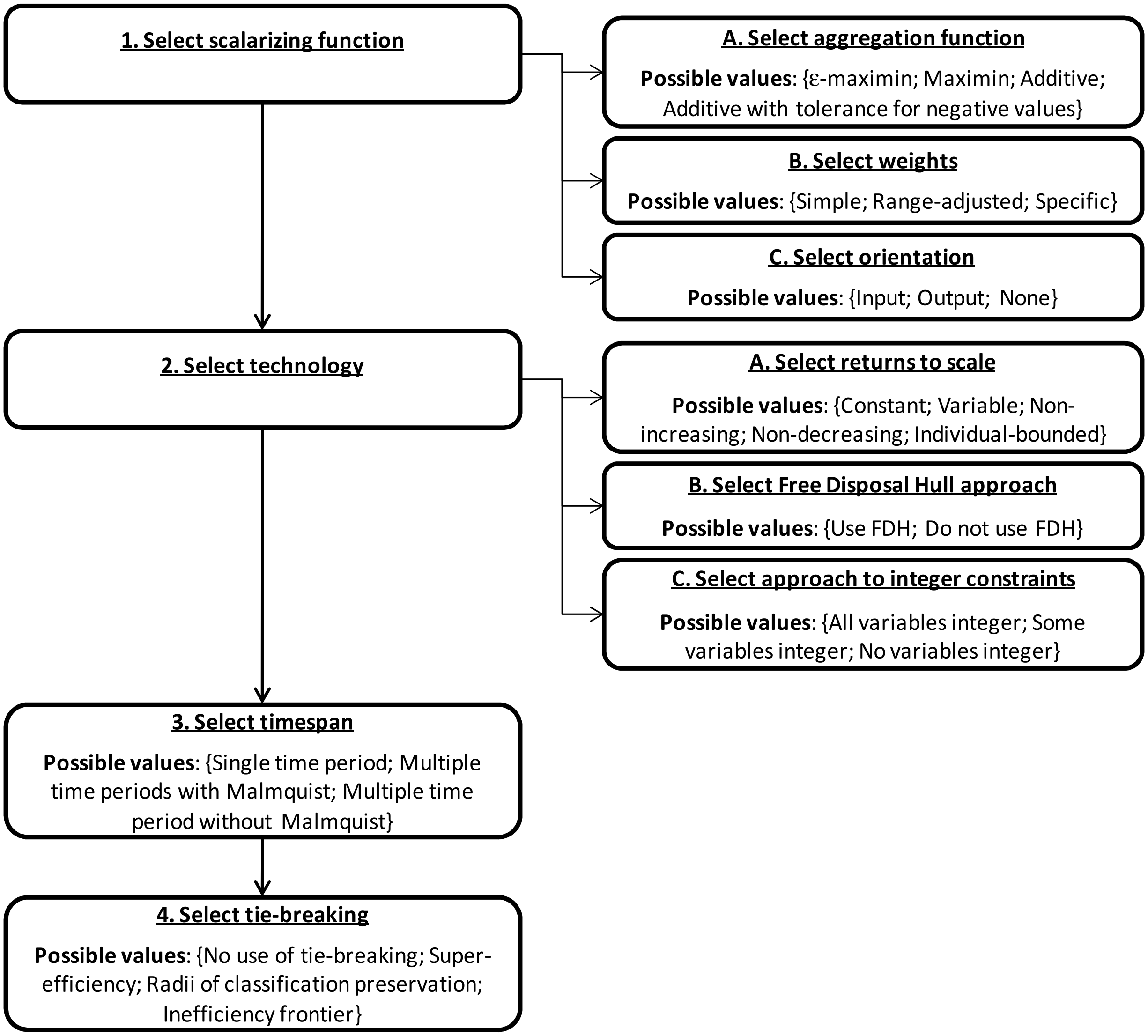

Figure 3.4 - Structure of a DEA model framework for airport benchmarking. Full details available in Appendix A.

3.3.2 Heuristics for Making Choices in the DEA Model Selection Framework

Choices in the DEA model selection framework should be based on the characteristics of the domain being modeled. The heuristics for making choices in the DEA model selection framework when benchmarking airport performance were developed by analyzing the characteristics of airport operations and their environment. A summary of the heuristics are presented in Table 3.3, with full details in Appendix A. Rules for translating the DEA model choices made using the heuristics are mapped to DEA model implementation parameters in Table 3.4.

Table 3.3 – Heuristics for airport benchmarking using the DEA model choice framework. Full details available in Appendix A.

Scalarizing function
Aggregation
Use either ε-maximin or additive. If the ignorance of slacks in the efficiency score is acceptable, then ε -maximin is the choice that reflects management's choices about the mix of inputs and/or outputs. Otherwise, use the additive function. In addition, if any parameters take on negative values then the additive function implemented in the additive model adjusted for negative data must be used.

Weights
Use specific weights unless evidence exists that range-adjusted weights are more appropriate.
Orientation
If the model requires orientation, then choose orientation to reflect which parameters are controllable by management.
Technology
Returns to scale
If modeling some version of labor and capital resources as inputs and passengers and aircraft movements as outputs, then use VRS. Otherwise, study the parameters to determine if VRS or CRS exist.
Free Disposal Hull
Unless compelling evidence that study results will be better accepted if only observed values are used for peer comparisons, do not use FDH.
Integer constraints
Use integer constraints for inputs and outputs with low magnitudes, such as runways.
Timespan

If modeling some version of labor and capital resources as inputs and passengers and aircraft movements as outputs over multiple time periods, then use a Malmquist index. For other domains, review if technology changes over time have occurred.

Tie breaking

If the study requires that all airports be fully ranked, use the tie-breaking function that provides the best intuitive interpretation; otherwise do not use a tie-breaking function.

Table 3.4 - Translation of heuristics to specific model choices

Element	Choice	Translation in modeling
Aggregation	ε-maximin	Use CCR or BCC with minimum bounds on weights, as described in sections 2.2.2.6.1 and 2.2.2.6.2.
	Additive	Use SBM/Russell measure, since these provide units invariant modeling options, as described in section 2.2.2.6.4.
	Additive with tolerance for negative data	Use the adaptation of the SBM/Russell measure model with tolerance for negative data, as described in section 2.2.2.6.11.
Weights	Specific weights	Use original model as specified.
Orientation	Input/output	If using an oriented model such as BCC, choose the input or output oriented version as appropriate.

Element	Choice	Translation in modeling
Technology	CRS	If the aggregation function is ε-maximin, then choose CCR, as described in section 2.2.2.6.1. If using some other model, ensure that no convexity constraint such as eλ=1 is present in the model.
	VRS	If the aggregation function is ε-maximin, then choose BCC, as described in section 2.2.2.6.2. If using some other model, ensure that a convexity constraint such as eλ=1 is present in the model.
Free Disposal Hull	Use FDH	Use the FDH implementation as described in section 2.2.2.6.5.
	No use of FDH	Use original model as specified.
Integer constraints	Some/all variables integer constraints	Use the implementation as described in section 2.2.2.6.12.
	No integer constraints	Use original model as specified.

Element	Choice	Translation in modeling
Timespan	Use Malmquist index	Use Malmquist index implementation as described in section 2.2.2.6.8.
	No Malmquist index	Use original model as specified.
Tie-breaking	Use tie-breaking	Use one of the implementations as described in sections 2.2.2.6.6, 2.2.2.6.9, or 2.2.2.6.10.
	No tie-breaking	Use original model as specified.

3.4 Step 4: Collect Data

Every benchmark of airport performance has a unique set of requirements for performance data, resulting in a review of available data being necessary for each benchmark. However, several key data sources that are applicable in many types of benchmarks can be identified.

This section provides an overview of those data sources. The sources provide raw data, with preprocessing being necessary in many cases to determine the aggregate or derived values necessary in the benchmark.

The data sources include:

- **Data on airline service**: Data on the traffic between airport pairs is available from the T100 database which is compiled from data collected by Office of Airline Information (OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b). It includes variables such as the frequency of service, the available seat capacity, and the number of passengers carried. The T100 database is a complete census of flights by U.S. and foreign carriers.

- **Airfare data**: Data on airfares is available from the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c)
- **Airport financial data**: The FAA's Compliance Activity Tracking System provides data on airport revenues and costs (Federal Aviation Administration 2010a)
- **Aircraft movement volume data**: The FAA's Air Traffic Activity System provides data on aircraft movements by type of aircraft (Federal Aviation Administration 2010)
- **Data on on-time performance**: On-time data is compiled from data collected by the OAI at the BTS (Bureau of Transportation Statistics 2010b). This data only encompasses U.S. carriers.
- **GDP data**: Data on GDP by metropolitan area is available from the U.S. government's Bureau of Economic Analysis (BEA) (Bureau of Economic Analysis, U.S. Department of Commerce 2010).
- **Population data**: Data on the population by metropolitan area is available from the U.S. Census Bureau (U.S. Census Bureau 2010b).

3.5 Step 5: Compute Benchmark Results

To compute the benchmark results, a software implementation of DEA models is used. This section describes commercially available DEA software as well as a DEA implementation for this dissertation.

Software for computing DEA scores is commercially available and academic DEA freeware can also be obtained. Table 3.5 provides an overview of a variety of available DEA software.

Table 3.5 - Overview of DEA software

Product	Pricing (as of June, 2010)
Banxia Software Frontier Analyst (Banxia 2010)	$285 to $5,850
DEA Frontier (DEA Frontier 2010)	$349 to $1,500
Performance Improvement Management DEA Software (Performance Improvement Management 2010)	$485 to $7,175
DEAP (Coelli 2010)	Freeware
DEA Solver Online (Kleine 2010)	Freeware

For the purposes of the DEA analysis described in this dissertation, these software products lack the ability for the user to customize modeling parameters to the degree necessary. To compute the DEA scores using the different models described in the DEA framework, the dissertation analysis included the implementation of several DEA models to allow for custom settings of the different model parameters.

Two different implementations were developed: One comprehensive implementation using Matlab, requiring the user to have access to a Matlab license; and a second, basic implementation in C++ which does not require the user to have access to any specialized software other than the CPLEX (IBM 2010) linear program solver and associated interfaces. Both implementations rely on CPLEX for solving the linear and mixed-integer programs. The architecture of the two implementations is provided in Figure 3.5 and the following two subsections describe the two implementations in more detail. The full code of the two implementations is provided in the Appendix B and Appendix C.

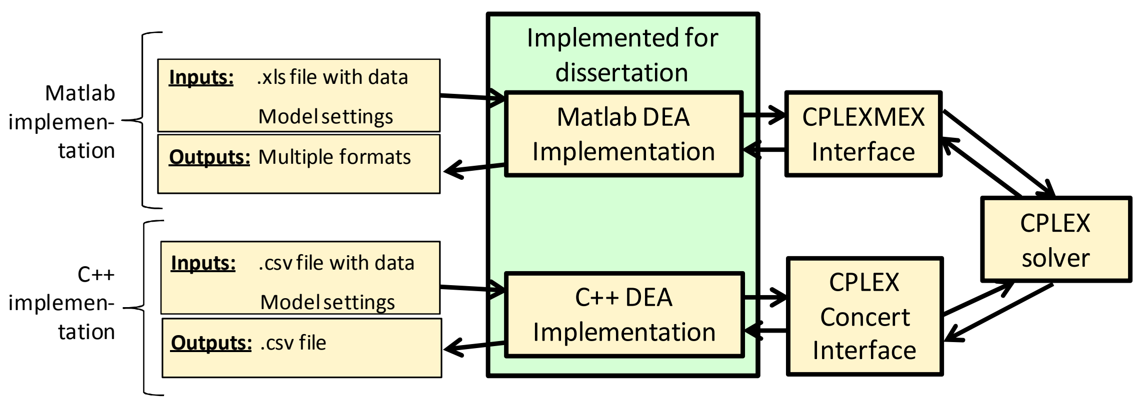

Figure 3.5 - Architecture of the two DEA implementations

3.5.1 Matlab Implementation

The Matlab implementation is provided in Appendix B and includes the implementation of the following DEA models:

- CCR (A. Charnes et al. 1978), discussed in section 2.2.2.6
- BCC (R. D. Banker et al. 1984), discussed in section 2.2.2.6.2
- SBM (Tone 2001), discussed in section 2.2.2.6.4
- Additive model adjusted for negative data (Zhu & Cook 2007, p. 76), discussed in section 2.2.2.6.11.

The implementation also allows for setting the orientation of the model, setting minimum weights, and setting integer constraints. The implementation interfaces with Matlab through the CPLEXMEX interface (Giorgetti 2010). See Appendix B for the full code.

3.5.2 C++ Implementation

The C++ implementation is provided in Appendix C and implements the input-oriented versions of the following models:

- CCR (A. Charnes et al. 1978), discussed in section 2.2.2.6
- BCC (R. D. Banker et al. 1984), discussed in section 2.2.2.6.2

- SBM (Tone 2001), discussed in section 2.2.2.6.4

The implementation interfaces with CPLEX through the CPLEX Concert technology. See Appendix C for the full code.

3.6 Step 6: Interpret Results and Derive Findings

With the benchmark results computed, the analysis shifts to interpreting the benchmark results to determine the controllable and uncontrollable factors which impact results. This encompasses the fourth step of the benchmarking process described in (McNair & Leibfried 1992).

This section describes three methods for investigating the characteristics of the benchmark results. All three of the methods, or a subset of them, may be applied for analyzing the benchmark results. The three methods are:

1. **Identifying the factors which impact the benchmark results**: This process serves to formulate hypotheses about the impact on the benchmark results of controllable factors (e.g. management practices) and uncontrollable factors (e.g. weather), and testing those hypotheses.

2. **Categorizing airports according to environmental factors**: This analysis serves to categorize the airports along two variables, one of which is the benchmark results and the second is some environmental variable (e.g. average yields). The purpose of this categorization is to identify groups of airports that meet a certain characteristic and to formulate an analysis of the implications of membership in those groups (e.g. what does it mean for airports that have poor benchmark performance and high average yields?).
3. **Investigation of individual airports' results**: This analysis serves to pinpoint an individual airport, usually one with outlier characteristics, and investigate which (potentially unique) factors impact that airport.

The following three subsections discuss each of these analytical approaches in turn.

3.6.1 Identifying Factors which Impact Benchmark Results

This step serves to identify factors which have an impact on benchmark results. The step involves formulating hypotheses about factors which impact results, and then testing those hypotheses.

The analysis begins with consulting existing research and writings about the aspect of airport performance that is being studied, as well as eliciting subject matter experts' views, with the purpose of compiling a set of hypotheses about factors which impact airport performance. The factors can include both factors which are considered controllable by airport management (e.g. the degree of outsourcing) as well as factors which are uncontrollable by management (e.g. weather conditions).

The statistical methods to be used for testing these hypotheses must be carefully chosen since the normality assumption may not hold true for DEA results. If the normality assumption does not hold, t-tests and regressions are not suitable methods for this step of the analysis. Rather, the methods must be those that do not require the normality assumption to hold true. Methods that have been used in past studies include:

- The Kruskal-Wallis test (Kruskal & Wallis 1952) (used in e.g. (Bazargan & Vasigh 2003)), which is based on the ranks of DMUs, and does not require the normality assumption to hold true.
- Tobit regressions (Tobin 1958) (used in e.g. (Gillen & Lall 1997)). Tobit regressions are suitable when the data represents a partial normal distribution, which may be the case with DEA benchmark results.

3.6.2 Categorizing Airports According to Environmental Factors

The categorization of airports according to environmental factors serves to classify sets of airports into groups which share similar characteristics. The purpose of this classification is to make it possible to analyze the nature and situation of the airports that fall within a particular group.

For instance, if the example begun earlier is continued, grouping the results of a benchmark of the level of air service according to the average yield provides groups such as "poor benchmark results and high yield", "poor benchmark results and low yield", etc. Each of these groups can then be analyzed to determine what the implications are of membership in a particular group.

To determine which factors to group airports by, the analyst should – as in the previous section about hypothesis formulation – use resources such as past research, writings, and interviews of subject matter experts.

The difference between the categorization of airports as described in this section, and identification of causal factors in the previous section, is that the categorization of airports is not focused on the causal effect on the benchmark results of the variables considered; rather the focus is to further refine the context of the results.

3.6.3 Investigating Individual Airports

The third type of investigation is focused on individual airports rather than the full set of airports included in the benchmark. Commonly, the purpose of this step is to understand the factors which cause an individual airport's position in the results. If the airport being investigated is one with particularly strong benchmark results, the purpose is to understand what enabled its strong performance; if the airport is one with poor results, the objective is to understand what caused its poor performance.

Two analytical strategies are available for this step:

1. Analyze the factors which contribute directly to the inputs and outputs used in the benchmark. For instance, if total operating cost is one input to the benchmark, analyze the subcategories of costs that make up the operating cost to identify if one such category stands out as higher or lower than the norm.
2. Review research and news sources that explain the particular environment in which the airport operates. For instance, if the volume of traffic has dropped sharply, news articles that describe the loss of airline hub service at that airport during the time period being studied would be a relevant piece of evidence.

4 Chapter 4: Results

The results comprise four components. First is an application of the DEA model selection framework and heuristics to the past airport benchmarks to assess their validity. Subsequently, the stakeholder model and model selection framework and heuristics are applied in three benchmark case studies, each presented in its own subsection:

1. A benchmark of the level of domestic air service to U.S. metropolitan areas
2. A benchmark of the level of capacity utilization for providing high levels of air service and moving large volumes of passengers at the U.S. OEP-35 airports
3. A re-design of a benchmark from the literature in which the original benchmark's premise is used as a starting point, and the stakeholder model and model selection framework and heuristics are applied to create a new benchmark

4.1 Assessment of Validity of Past Benchmarks: Application of the DEA Model Selection Framework and Heuristics to Past Airport Benchmarks

The purpose of this section is to apply the DEA modeling framework and heuristics on 13 past airport benchmarks to identify if discrepancies exist between what the framework and heuristics prescribe and what the study authors chose in their models. Such discrepancies have an impact on the validity of the findings from these studies. Discrepancies also reinforce the need for a more systematic approach to DEA model selection in future studies.

The first subsection applies the framework and heuristics to the past studies of airport performance. The following subsection analyzes the findings from applying the framework and heuristics and the last subsection analyzes the implications of the findings.

4.1.1 Results of Framework and Heuristics Application

The past studies of airport performance using DEA were reviewed using the DEA framework and heuristics. Each study was assessed in three steps:

- **Study choice**: Each study was analyzed to determine which choice the authors made for each element in the DEA framework.
- **Analysis**: An analysis of the study was made for each element in the framework based on the objective of the study, its domain, the inputs and outputs used, and the timespan of the analysis.
- **Recommendation according to framework**: Using the heuristics from 0, the recommended choice was determined for each element in the framework.

Using this information, it was possible to compare the choices made by the study authors with the choices recommended by the heuristics. The results of this analysis are presented in Table 4.1 and the underlying analysis of each individual study is presented subsequent to the table.

Table 4.1 - Results of application of DEA framework and heuristics to past academic airport benchmarks

Study	Do the Study's Choices Agree with the Framework and Heuristics?								Level of agreement with framework and heuristics
	Scalarizing function			Technology					
	Aggregation	Weights	Orientation	Returns to scale	FDH	Integer constr.	Time-span	Tie-breaking	
Size Versus Efficiency: A Case Study of US Commercial Airports Airports (Bazargan & Vasigh 2003)	✓	✓	✓	✗	✓	✗	✗	✗	50%
Relative Efficiency of European Airports (Pels et al. 2001)	✗	✓	✓	✓	✓	✗	✗	✓	63%
Developing Measures of Airport Productivity and Performance - Airside study (Gillen & Lall 1997)	✗	✓	✓	✗	✓	✗	✗	✓	50%
Developing Measures of Airport Productivity and Performance - Terminal study (Gillen & Lall 1997)	✗	✓	✓	✓	✓	✗	✗	✓	63%
Performance Based Clustering for Benchmarking of US	✗	✓	✓	✗	✓	✗	✗	✓	50%

Study	Do the Study's Choices Agree with the Framework and Heuristics?								Level of agreement with framework and heuristics
	Scalarizing function			Technology					
	Aggre-gation	Weights	Orien-tation	Returns to scale	FDH	Integer constr.	Time-span	Tie-breaking	
Airports (Sarkis & Talluri 2004)									
An application of DEA to measure the efficiency of Spanish airports prior to privatization (Martín & Román 2001)	✓	✓	✓	✓	✓	✓	✓	✓	100%
Measuring Airport Quality from the Airlines' Viewpoint (Adler & Berechman 2001)	✗	✓	✓	✗	✓	✓	✓	✗	63%
An analysis of the operational efficiency of major airports in the United States (Sarkis 2000)	✗	✓	✓	✗	✓	✗	✗	✓	50%
Managerial Efficiency of Brazilian Airports (Pacheco & Fernandes 2003)	✓	✓	✓	✓	✓	✓	✓	✓	100%
The performance of BAA before and after privatization (Parker 1999)	✓	✓	✓	✗	✓	✓	✗	✓	75%
Total factor productivity and	✗	✓	✓	✓	✓	✓	✓	✓	88%

Study	Do the Study's Choices Agree with the Framework and Heuristics?								Level of agreement with framework and heuristics
	Scalarizing function			Technology					
	Aggre-gation	Weights	Orien-tation	Returns to scale	FDH	Integer constr.	Time-span	Tie-breaking	
efficiency of Australian airports (Abbott & Wu 2002)									
Airports in Argentina: Technical efficiency in the context of an economic crisis (Barros 2008)	✗	✓	✓	✗	✓	✓	✗	✓	63%
Performance evaluation of Italian airports: A data envelopment analysis (Barros & Dieke 2007)	✓	✓	✓	✓	✓	✓	✓	✓	100%

Level of agreement with framework and heuristics	38%	100%	100%	46%	100%	54%	38%	85%

The following tables present the analysis that underlies the conclusions presented in Table 4.1.

Table 4.2 - Application of DEA framework and heuristics to past airport benchmark studies

		Size Versus Efficiency: A Case Study of US Commercial Airports (Bazargan & Vasigh 2003)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	ε-maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✓
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	N/A	N/A	N/A	✓
Technology	**Returns to scale**	Constant	These inputs and outputs should be modeled with variable returns to scale.	VRS	✗
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of runways; Number of jetways.	✗
Timespan		Multiple time periods without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

		Relative Efficiency of European Airports (Pels et al. 2001)			
			DEA Framework and heuristics		Agreement between the author's choices and the framework and heuristics
		Study choice	Analysis	Recommendation	
Scalarizing function	Aggregation	Maximin	All parameters should be considered in the aggregation study.	ℇ-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Unknown	Recommend output orientation for this study since the inputs are largely fixed assets which cannot easily be removed. Promoting increased traffic, etc., is more controllable by management.	Output oriented	✓
Technology	Returns to scale	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of remote stands; number of terminal parking positions; number of check-in desks; number of baggage claims.	✗
Timespan		Multiple time periods without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗

Tie-breaking				

		Developing Measures of Airport Productivity and Performance - Airside study (Gillen & Lall 1997)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	Maximin	All parameters should be considered in the aggregation study.	Ɛ-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	N/A	N/A	N/A	✓
Technology	**Returns to scale**	Constant	These inputs and outputs should be modeled with variable returns to scale.	VRS	✗
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of runways	✗
Timespan		Multiple time periods grouped into one without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

Developing Measures of Airport Productivity and Performance - Terminal study (Gillen & Lall 1997)		
Study choice	**DEA Framework and heuristics**	**Agreement between the**

			Analysis	Recommendation	author's choices and the framework and heuristics
Scalarizing function	Aggregation	Maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Output oriented	Either output or input oriented could be used since both inputs (number of employees) and outputs (traffice volumes) include parameters that management can control/influence.	Indifferent	✓
Technology	Returns to scale	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of runways; Number of gates; Number of baggage collection belts	✗
Timespan		Multiple time periods grouped into one without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

Performance Based Clustering for Benchmarking of US Airports (Sarkis & Talluri 2004)

			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	Maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	N/A	N/A	N/A	✓
Technology	**Returns to scale**	Constant	These inputs and outputs should be modeled with variable returns to scale.	VRS	✗
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of runways; Number of gates	✗
Timespan		Multiple time periods without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

An application of DEA to measure the efficiency of Spanish airports prior to privatization (Martín & Román 2001)			
	DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
Study choice	**Analysis**	**Recommendation**	

Scalarizing function	**Aggregation**	ε-maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✓
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Output oriented	The inputs are fixed and cannot be controlled by management.	Output oriented	✓
Technology	**Returns to scale**	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Single	N/A	N/A	✓
Tie-breaking					

		Measuring Airport Quality from the Airlines' Viewpoint (Adler & Berechman 2001)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	Maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Input oriented	The inputs such as airport charges are controllable by airport management.	Input oriented	✓

	Returns to scale	Variable	This cannot be evaluated since the authors mix metrics that increase with scale (e.g. number of runways) with metrics that are independent of scale (e.g. landing fee/movement). This is a violation of DEA modeling rules.	N/A	✗
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
Technology	**Integer constraints**	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Single	N/A	N/A	✓
Tie-breaking					

		An analysis of the operational efficiency of major airports in the United States (Sarkis 2000)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
	Aggregation	Maximin	All parameters should be considered in the aggregation study.	Ɛ-maximin or additive	✗
Scalarizing function	**Weights**	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓

	Orientation	Unknown	Either output or input oriented could be used since both inputs (number of employees) and outputs (traffice volumes) include parameters that management can control/influence.	Indifferent	✓
Technology	**Returns to scale**	Mixed - uses different models.	VRS exist for these inputs and outputs.	VRS	✗
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	Some inputs require integer constraints.	Inputs that require integer constraints include: Number of runways; Number of gates	✗
Timespan		Multiple time periods without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

		Managerial Efficiency of Brazilian Airports (Pacheco & Fernandes 2003)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	ε-maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✓
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓

	Orientation	Input oriented	The inputs such as costs and number of employees are controllable by airport management.	Input oriented	✓
Technology	**Returns to scale**	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓
	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Single	N/A	N/A	✓
Tie-breaking					

		The performance of BAA before and after privatization (Parker 1999)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	e-maximin or maximin (unknown which one was used)	All parameters should be considered in the aggregation study.	e-maximin or additive	✓
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Unknown	The inputs such as costs and number of employees are controllable by airport management.	Input oriented	✓
Technology	**Returns to scale**	Mixed - uses different models.	VRS exist for these inputs and outputs.	VRS	✗

	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		One airport is benchmarked against its own performance over time.	This modeling is problematic since technology changes occurred during the time period. However, since the same entity is benchmarked against its own performance, it is not possible to compute a Malmquist index.	N/A	✗
Tie-breaking					

		Total factor productivity and efficiency of Australian airports (Abbott & Wu 2002)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	Maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Input oriented	Some inputs such as the number of employees are controllable by airport management.	Input oriented	✓
Technology	**Returns to scale**	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓

	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Multiple time periods with Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✓
Tie-breaking					

		Airports in Argentina: Technical efficiency in the context of an economic crisis (Barros 2008)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
Scalarizing function	**Aggregation**	Maximin	All parameters should be considered in the aggregation study.	ε-maximin or additive	✗
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
	Orientation	Output oriented	Either output or input oriented could be used since both inputs (number of employees) and outputs (traffic volumes) include parameters that management can control/influence.	Indifferent	✓
Technology	**Returns to scale**	Mixed - uses different models.	These inputs and outputs should be modeled with variable returns to scale.	VRS	✗

	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Multiple time periods without Malmquist	These inputs and outputs should be modeled with a Malmquist index since technology change is observed over time.	Multiple time periods with Malmquist	✗
Tie-breaking					

		Performance evaluation of Italian airports: A data envelopment analysis (Barros & Dieke 2007)			
			DEA Framework and heuristics		**Agreement between the author's choices and the framework and heuristics**
		Study choice	**Analysis**	**Recommendation**	
	Aggregation	Not specified	All parameters should be considered in the aggregation study.	Ɛ-maximin or additive	✓
	Weights	Specific	Assume that airports are making tradeoffs of their own.	Specific	✓
Scalarizing function	**Orientation**	Unknown	Either output or input oriented could be used since both inputs (e.g. labor costs) and outputs (aeronautical revenue) include parameters that management can control/influence.	Indifferent	✓
Technology	**Returns to scale**	Variable	These inputs and outputs should be modeled with variable returns to scale.	VRS	✓

	FDH	No	No reason to only compare to observed input/output combinations.	No	✓
	Integer constraints	None	No inputs or outputs require integer constraints.	None	✓
Timespan		Single	N/A	N/A	✓
Tie-breaking					

These findings show that past studies have had no issues with selection of model weights, determination of model orientation, or the use of FDH. However, discrepancies do exist in the selection of aggregation function, choice of returns to scale, the use of integrality constraints, and the application of Malmquist indices. In the table, a study was not considered to have any discrepancy of the author's choice was unknown. In total, three of the 13 studies exhibited no discrepancies relative to the framework and heuristics.

4.1.2 Areas of Discrepancy between Study Choices and Heuristics

This section discusses those areas where discrepancies exist between a number of studies and the recommendations by the heuristics.

4.1.2.1 Aggregation function

Eight studies use the maximin aggregation function. This has the implication that some inputs and outputs may be ignored in the calculation of the objective function value. The heuristics described in 0 indicate that only ε-maximin or additive aggregation functions should be used in airport benchmarking. The rationale for this is that the inclusion of the inputs and outputs in the analysis must be based on a determination about which parameters are important to the analysis and as a result all parameters should be considered to some degree in the determination of each airport's DEA score (R. G. Dyson et al. 2001, p. 253).

The impact of the use of the maximin aggregation function is that the studies which have used this aggregation function are likely to include many DMUs which have achieved their efficiency score by assigning all weights to only one input and/or one output. This could mean for instance that an airport has been considered fully efficient by achieving a strong ratio of passengers to runways, while all other resources such as labor costs are over-consumed in relation to the levels of passengers and aircraft movement.

4.1.2.2 Returns to Scale

The analysis in Table 4.2 shows that seven studies have used the assumption of constant returns to scale in their modeling of airport performance when research in fact indicates that the domain being modeled reflects VRS. The fact that the modeling assumptions do not reflect real-life conditions means that the results of the analysis may not be valid.

VRS models have a convexity constraint that is not present in CRS models (William W. Cooper et al. 2006, p. 87). This added constraint implies that the feasible region for VRS models such as BCC is smaller than that of CRS models like CCR (William W. Cooper et al. 2006, p. 88). The implication is that studies that have applied CRS models may have deemed airports inefficient that are, in fact, fully efficient.

4.1.2.3 Integer constraints

Although none of the 13 studies examined applied integrality constraints on inputs or outputs, only six of these studies used parameters which require integrality constraints. The parameters which require integrality constraints according to the heuristics are:

- Number of runways (5 studies use this measure)
- Number of jetways/gates/terminal parking positions (5 studies) (these inputs are not identical, but are very close)
- Number of baggage collection belts (2 studies)
- Number of remote parking stands (1 study)
- Number of check-in desks (1 study)

The heuristics indicate that measures with low magnitude should be subjected to integrality constraints. The number of runways has a particularly low magnitude, indicating that the integrality constraints are particularly important for this parameter.

Applying integrality constraints will shrink the feasible region which means that efficiency scores in models with integrality constraints will be greater than or equal to those computed for the same models without integrality constraints. The implication for existing studies which did not apply integrality constraints is that it is possible that some DMUs which should have been deemed fully efficient were, in fact, rated as inefficient.

4.1.2.4 Timespan

Eight among the 13 studies examined included analyses across multiple time periods without computation of a Malmquist index in spite of research evidence that technology changes had in fact occurred during the time period being analyzed. The implication of not using the assumption of technology changes during the course of time is the same as when modeling CRS when reality reflects VRS; the model does not effectively mirror reality and the results derived from these studies may be questioned. In these seven studies, efficiency scores may have been over or under-estimated.

4.1.3 Implications of the Gaps between Study Choices and Heuristics

The analysis showed four areas of discrepancy:

1. Use of aggregation function which permits some inputs and outputs to be ignored
2. Inappropriate use of constant returns to scale in model
3. Lack of use of integrality constraints for some model inputs
4. Lack of computation of a Malmquist index when analyzing performance across multiple time periods

The implications of these findings are twofold:

First, the conclusions drawn from these studies may not hold. This extends not only to questioning of the efficiency scores determined for the airports included in the studies but also to questioning of the analyses done of exogenous factors impacting these efficiency scores. For instance, the finding that small hubs are more efficient than large hubs (Bazargan & Vasigh 2003) or the finding that terminal efficiency is greater if the airport is operated under compensatory funding rules[2] (Pels et al. 2001) may not have held if the modeling assumptions were different.

Second, the implication for future airport benchmarking studies is that a need exists for systematic application of a framework and heuristics for DEA model selection. Applying the model selection framework and heuristics from Appendix A will result in future analyses which will withstand methodological scrutiny.

[2] Compensatory funding means that air carriers pay predetermined usage fees to the airport organization. If traffic volumes drop, the total revenue for the airport drops. Compensatory funding is contrasted with residual funding, which means that the shortfall after non-aeronautical revenues have been applied toward the cost of running the airport is divided among the air carriers using the airport (Doganis 1992) (p. 192).

4.2 Case Study 1: Benchmarking the Level of Domestic Air Service to U.S. Metropolitan Areas

With major airports in the U.S. operating under profit-neutral financial regulations (Carney & Mew 2003, p. 230), as “public utilities,” they play an important role in shaping the national airline transportation system. In service to multiple regional stakeholders (Schaar & Sherry 2010), airport authorities incentivize the type and quantity of airline transportation service provided (Belobaba et al. 2009, pp. 168-175), (A. Graham 2003, p. 189).

This case study presents the results of an analysis of the degree to which the level of domestic airline service has been maximized in relation to the size of regional economies and populations. A DEA benchmark was used to determine “best-in-class” in terms of frequency and number of destinations served based on the size of the regional economy and population.

The results indicated that 20 of the top 29 metropolitan areas have high levels of air service but that nine regions exhibit gaps in their level of service relative to the size of their population and regional economy

These results have implications for strategic planning on a national scale, airport improvement funding, and regional planning. Whereas flight delays are indicative of insufficient capacity, the more important question is whether the existing airport resources are being used most efficiently.

This case study is organized as follows: Section 4.2.1 reviews the airport stakeholders and their goals related to the level of air service. Section 4.2.2 discusses the study methodology, including the means for selecting performance parameters and the benchmarking model used. Section 4.2.3 reviews the study results. Section 4.2.4 presents the conclusions.

4.2.1 The Airport's Stakeholders and Their Goals

The analysis of airport stakeholders in section 2.1.3 found that the stakeholders' goals for the airport were based in part on factors wholly within the control of airport management (the "airport organizational boundary"), but also on factors that were only partly within the control of management, or entirely outside management's control.

The goal of "maximizing the number of destinations served and frequency of those services" emerged from the analysis as common to stakeholder groups such as local businesses, residents, the local government, and the airport organization itself. It is an example of a goal that is not fully within the control of airport management since airlines determine where to add or reduce service.

The goal reflects a "symbiotic" relationship between a region's economy and the local air service, where air service stimulates economic growth (Button & Stough 2000) and growth in a region's economy drives increased demand for air travel (Intergovernmental Panel on Climate Change 2000).

The stakeholders who are concerned with this goal have a need for evaluating the degree to which it is being achieved in U.S. metropolitan areas. Local governments and airport authorities must understand if their region is currently well served by airlines or if added effort is necessary to attract additional air service. If a shortfall exists in the degree to which the goal is being met, they must gain insight about what is causing the performance gap. Conversely, a region's residents and business community must understand if their needs are being met by the airport(s) in their region, or if they should demand more from their local government and airport authority in terms of attracting new air service to their community.

A comparative benchmark is a means to evaluate this goal. The benchmark allows for a normalized comparison across major U.S. metropolitan areas and gives stakeholders an understanding of which areas are not currently well served and can also provide insight into the causes of any performance gaps.

4.2.2 Methodology

This section discusses the study methodology. It provides the basis for the selection of performance parameters and discusses the choice of model for benchmarking. It also describes the data sources and pre-processing as well as the method used for computing benchmark scores. Finally, it presents the method for sensitivity analysis of the results.

4.2.2.1 Scope of Analysis

The study reviews the levels of air service in metropolitan areas. Some metropolitan areas include multiple airports (e.g. the Boston metropolitan area, with Boston-Logan, Providence, and Manchester airports) and other areas are served by a single airport (e.g. Atlanta). Table 4.3 shows the airports included in the study, organized by metropolitan area. A full description of the methodology for determining metropolitan areas and mapping airports to those areas is provided in section 4.2.2.4.

Table 4.3 - Airports included in study

Metropolitan Area	Airport Name	Airport Code
Atlanta	Hartsfield - Jackson Atlanta International	ATL
Boston	General Edward Lawrence Logan International	BOS
	Manchester	MHT
	Theodore Francis Green State	PVD
Charlotte	Charlotte/Douglas International	CLT
Chicago	Chicago Midway International	MDW
	Chicago O'Hare International	ORD
Cincinnati	Cincinnati/Northern Kentucky International	CVG
	James M Cox Dayton International	DAY
Cleveland	Cleveland-Hopkins International	CLE
Dallas	Dallas Love Field	DAL
	Dallas/Fort Worth International	DFW
Denver	Denver International	DEN
Detroit	Detroit Metropolitan Wayne County	DTW
Honolulu	Honolulu International	HNL
Houston	William P Hobby	HOU
	George Bush Intercontinental/Houston	IAH

Metropolitan Area	Airport Name	Airport Code
Las Vegas	McCarran International	LAS
Los Angeles	Los Angeles International	LAX
	Ontario International	ONT
	Bob Hope	BUR
	John Wayne Airport-Orange County	SNA
	Long Beach /Daugherty Field/	LGB
Memphis	Memphis International	MEM
Miami	Fort Lauderdale/Hollywood International	FLL
	Miami International	MIA
	Palm Beach International	PBI
Minneapolis	Minneapolis-St Paul International/Wold-Chamberlain	MSP
New York	John F Kennedy International	JFK
	La Guardia	LGA
	Newark Liberty International	EWR
	Long Island MacArthur	ISP
Orlando	Orlando International	MCO
Philadelphia	Philadelphia International	PHL
Phoenix	Phoenix Sky Harbor International	PHX
Pittsburgh	Pittsburgh International	PIT

Metropolitan Area	Airport Name	Airport Code
Portland	Portland International	PDX
Salt Lake City	Salt Lake City International	SLC
San Diego	San Diego International	SAN
San Francisco	San Francisco International	SFO
	Norman Y. Mineta San Jose International	SJC
	Metropolitan Oakland International	OAK
Seattle	Seattle-Tacoma International	SEA
St. Louis	Lambert-St Louis International	STL
Tampa	Tampa International	TPA
Washington-Baltimore	Ronald Reagan Washington National	DCA
	Washington Dulles International	IAD
	Baltimore/Washington International Thurgood Marshall	BWI

4.2.2.2 Selection of Model Parameters

Section 4.2.1 described one of the airport's goals as being to "maximize the number of destinations served and frequency of those services". To conduct a benchmark of the level to which this goal is achieved in each metropolitan area, the goal is translated into performance parameters that can be measured.

4.2.2.2.1 Measuring the Level of Air Service

The goal includes maximizing both the number of destinations served, as well as the frequency of those services. Two performance metrics are proposed in order to gauge the level to which this goal is achieved:

The first measure is the number of non-hub destinations served nonstop from any airport in the metropolitan area. The number of non-hub destinations served is defined in greater detail in a subsequent section, and refers to the number of destinations served other than those destinations that are determined to be hubs. This measure maps directly to the goal. Destinations which were served only on an occasional basis should not be considered and a lower bound of service at least once per week is imposed.

The second measure is the average daily frequency of service to the top domestic hubs (the definition of top domestic hubs is treated in section 4.2.2.4). This measure addresses the goal in two ways:

- It gives an indication of the level of frequency of service across a set of key routes
- It is a measure of the level of ease with which a large number of destinations can be reached through a single connection

These two measures reflect the two factors that impact total trip time, as discussed by (Belobaba et al. 2009, pp. 58-59). Total trip time involves both the time on board the aircraft as well as "schedule displacement," with the latter being the amount of time that passes between a passenger's desired departure time and the time when a flight is available. The number of destinations served nonstop will contribute toward minimizing the time on board the aircraft, and a high frequency of flights will minimize the schedule displacement.

4.2.2.2.2 Normalizing the Level of Air Service

Demand for air services come from a region's individual residents and businesses, as well as from business and tourist visitors to the region. Although some airports' passenger traffic is made up more heavily of connecting traffic and other airports' traffic to a greater degree consists of origin and destination (O&D) passengers, the number of individuals that reside in the region and the level of business activity are key drivers of the level of demand for air service, as shown in Figure 4.1 and Figure 4.2.

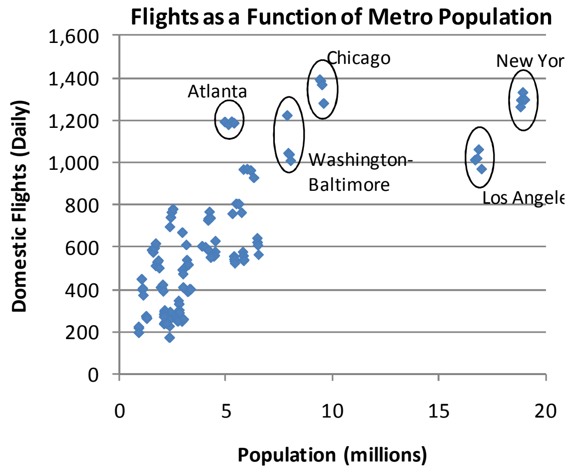

Figure 4.1 - Relationship between metropolitan area population and the number of domestic flights for the metro areas in Table 4.3, 2005-2008

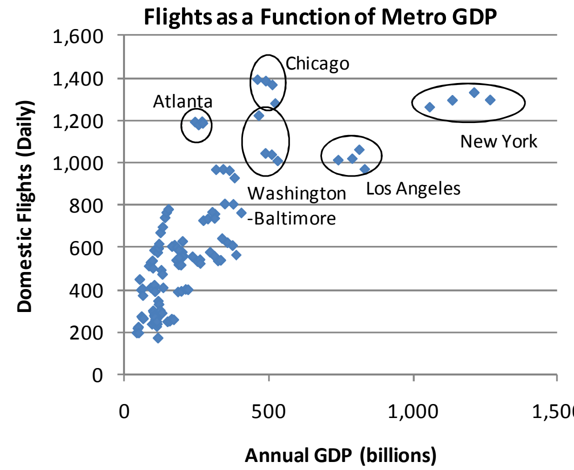

Figure 4.2 – Relationship between metropolitan area GDP and the number of domestic flights for the metro areas in Table 4.3, 2005-2008

The relationship between the population and the regional Gross Domestic Product (GDP) was tested and showed a high degree of correlation, with a Pearson coefficient of 0.78 ($p < 2.2 * 10^{-16}$). This correlation indicates that as the population goes up, so does the regional GDP, and vice versa. The relationship between the two parameters can be expressed as the GDP per capita, where the regional GDP is divided by the population. In spite of the high degree of correlation between the two parameters, a range of values for the GDP per capita exist between different metropolitan areas, as shown in Figure 4.3.

To account for the impact of both population and GDP on the level of flights in metropolitan areas, and to address the goals of both the region's population as well as its businesses, the benchmark data for the levels of air service should be normalized to account for the region's population and its regional GDP.

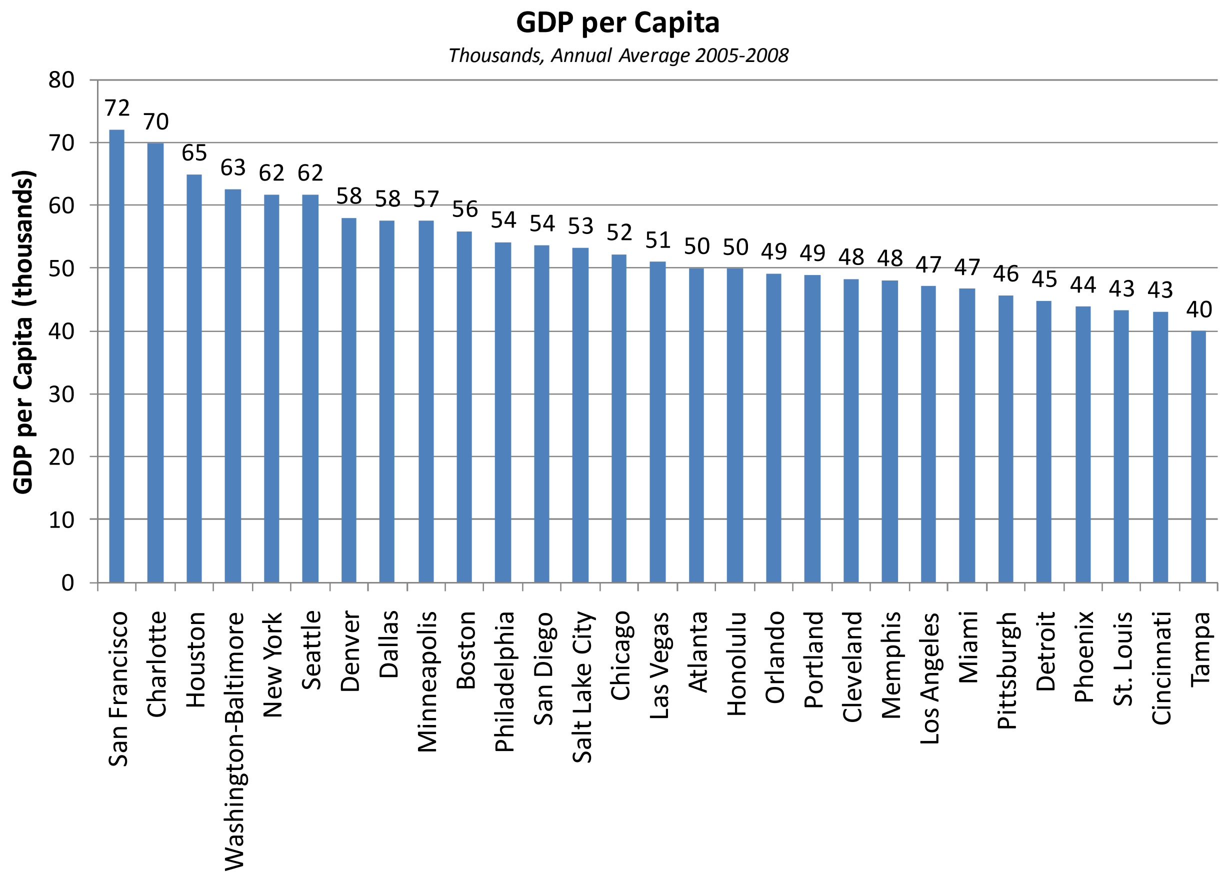

Figure 4.3 – Annual GDP per capita (thousands of US$), 2005-2008

4.2.2.2.3 Summary of Model Parameters

The measures of the level of air service and the parameters used to normalize them are combined in this conceptual ratio:

(destinations served nonstop, frequency of service to hubs) : (population, GDP)

The metropolitan areas with the highest number of destinations served and the highest frequency in relation to their population and GDP will be considered to have the highest relative level of air service.

4.2.2.3 Choice of Benchmark Model

The parameters for the model are the number of nonstop non-hub destinations served and the average daily frequency of service to the top domestic hubs, normalized by regional population and GDP. This model can conceptually be expressed as the ratio (destinations served, frequency) : (population, GDP). The units of measure for these metrics are airports, daily flights, people, and US$, respectively. Combining these metrics into a comparative benchmark is a case where the analysis combines multiple parameters of different units, and where the production or utility function is unknown. In this scenario, DEA is an appropriate method for calculating the composite benchmark scores, as shown in the benchmarking methodology decision tree in section 2.2.2.1.

The results of the application of the DEA framework and heuristics from Appendix A to determine a model for this analysis are now presented.

- **Aggregation**: The heuristics specify that either an ε-maximin function or an additive function should be used. The additive function should be used only if a motivation exists for why the current proportional mix of inputs or outputs (depending on the orientation chosen) is irrelevant and can be changed. Otherwise, the ε-maximin function should be chosen. In this study, no evidence exists to suggest that the proportional mix of input or outputs can be changed between different metropolitan areas. As a result, the ε-maximin function is chosen.
- **Weights**: Since tradeoffs between the two outputs will be different between metropolitan areas, specific weights should be used according to the heuristics.
- **Orientation**: The heuristics state that the model orientation should be determined based on which factors are considered the most controllable by management. In this analysis, the population and GDP inputs cannot be controlled by airport management, but although they are not directly controllable, the output measures of destinations

served and frequency can be influenced by airport management and local governments. This influence can come through providing air carriers with market research data as well as with financial incentives and marketing support for providing service to the airport (A. Graham 2003, p. 189). This determines this analysis as output-oriented.

- **Returns to scale**: The framework specifies a choice between constant returns to scale (CRS) and variable returns to scale (VRS). The outputs in this model can both be assumed to reflect VRS: First, the number of new destinations which are feasible to serve decreases as the number of already served destinations increases, since only a finite number of metropolitan areas exist where the local market provides sufficient demand to warrant nonstop service. Second, the potential for increased frequency of nonstop service to hubs declines as the level of existing frequency and airport congestion increases; in a hypothetical case, rather than providing service on a market every 5 minutes with a 50-seat aircraft, providing service every 10 minutes with a 100-seat aircraft would become necessary as airport capacity runs out (as utilization of airport capacity approaches its physical limit, policy/regulation changes may be necessary to incent airlines to fly larger aircraft (Donohue et al. 2008, pp. 115-116)).

- **FDH**: The Free Disposal Hull should be applied only if some reason exists why comparison only to observed combinations of inputs and outputs should be made, but no such reason exists in this analysis.
- **Integer constraints**: Integer constraints should be applied in cases where input or outputs are indivisible into fractions and of low magnitude, and if large errors in the results would be introduced if these inputs or outputs were assumed to have decimal values. The parameter with integer constraints and the lowest magnitude in this study is the number of non-hub destinations served nonstop, but with a median value of 88 for the years studied, this parameter's magnitude remains sufficiently high that no integrality constraints are necessary in the model.
- **Timespan**: If any key technology changes have occurred during the timespan being studied that would impact the ability of DMUs to achieve strong performance, then a Malmquist index method should be used. If not, the performance for each year can simply be analyzed independently. In the present analysis, technology changes would involve the introduction of something which made it feasible for air carriers to serve more destinations than before, or something which

allowed for increased frequency of service. From a technology point of view, this would involve the introduction of new aircraft types with highly different performance characteristics in terms of for instance fuel consumption, crew requirements, or number of seats. No new aircraft models for domestic use entered into service during the 2005-2008 period from Boeing (The Boeing Company 2010), Airbus[3] (Airbus S.A.S. 2010), Bombardier (Bombardier 2010), or Embraer (Embraer 2010). As a result of no major changes occurring in this time period, no Malmquist index calculation is necessary.

- **Tie-breaking**: The heuristics prescribe that a tie-breaking function be used only if a reason exists why all areas must be fully ranked. No such reason is present and accordingly, no tie-breaking function is used.

Table 4.4 summarizes the modeling assumptions for this analysis.

---

[3] The Airbus A380 was in fact first delivered in 2007, but this aircraft was not used for US domestic service

Table 4.4 - DEA model parameter choices

Scalarizing function			Technology			Timespan	Tie-breaking
Aggre-gation	Weights	Orient-ation	Returns to scale	FDH	Integrality		
ε-maximin	Specific weights	Output oriented	VRS	No use of FDH	No integrality constraints	No use of Malmquist index; simply one analysis per year	None

These modeling assumptions are represented in the output-oriented BCC (R. D. Banker et al. 1984) algorithm with minimum weight constraints, which was used in this analysis. This model has the following dual problem formulation and is discussed in greater detail in section 2.2.2.6.2:

$$\max(\phi_a, \lambda) = \phi_a + \varepsilon(s^+ + s^-)$$

$$\text{Subject to} \quad \begin{aligned} &\phi_a y_a - Y\lambda + s^+ = 0 \\ &X\lambda + s^- = x_a \\ &e\lambda = 1 \\ &\lambda \geq 0, s^+ \geq 0, s^- \geq 0 \end{aligned}$$

The DEA scores were computed by the BCC algorithm implementation in Matlab, as described in section 3.5.1. In the analysis, the infinitesimal constant ε

was set to 1.0 * E-6. A further discussion of the choice of this value is provided in section 4.2.3.5.1.

4.2.2.4 Data Collection and Pre-Processing

This section describes the means of obtaining and preparing the benchmark data for the analysis.

4.2.2.4.1 Determination of Metro Areas

The scope of the analysis was to include the metropolitan areas which have at least one of the OEP-35 airports listed in Table 0.1, and expand the study to include any other commercial airports that also service those metropolitan areas from within a given distance. In a second step, if any of the non-OEP-35 airports were located in a different nearby, second metropolitan area, then that second metropolitan area was merged with the first in order to capture the region's full population and GDP.

The definitions of "metropolitan areas" follow those of the U.S. government's Office of Management and Budget (OMB). The OMB defines "Metropolitan Statistical Areas" (MSAs) based on data from the Census Bureau (Office of Management and Budget 2010).

In their discussion of Multi-Airport Systems, (Neufville & Odoni 2003, p. 133) propose that studies only include airports that serve at least 1 million passengers per year. That limit is used in this analysis and only the 55 non-OEP-35 airports which met that criterion for at least one year between 2005 and 2008 were included for consideration.

A distance limit of 70 road miles from the city center of the main metropolitan area was used to determine which among the non-OEP-35 airports to include in the study, resulting in a final list of 13 additional airports, as shown in Table 4.5.

Table 4.5 - Non-OEP-35 airports added to the study

Airport Name	Airport Code
Bob Hope	BUR
Dallas Love Field	DAL
James M Cox Dayton International	DAY
William P Hobby	HOU
Long Island MacArthur	ISP
Long Beach /Daugherty Field/	LGB
Manchester	MHT
Metropolitan Oakland International	OAK
Ontario International	ONT
Palm Beach International	PBI
Theodore Francis Green State	PVD
Norman Y. Mineta San Jose International	SJC
John Wayne Airport-Orange County	SNA

With the addition of the 13 airports to the metropolitan areas, the locations of those airports which were situated in another, nearby metropolitan area were merged with the original metropolitan areas to accurately reflect the area's total population and GDP. Those areas were:

- The Manchester-Nashua, NH, MSA and the Providence-New Bedford-Fall River, RI-MA, MSA which were added to the Boston metropolitan area.
- The Dayton, OH, MSA which was added to the Cincinnati metropolitan area.
- The Riverside-San Bernardino-Ontario, CA, MSA which was added to the Los Angeles metropolitan area.
- The San Jose-Sunnyvale-Santa Clara, CA, MSA which was added to the San Francisco metropolitan area.

Finally, the Washington, DC, and Baltimore, MD, metropolitan areas were merged into one single area since the three airports serving the two cities are located within 61 miles of the two city centers.

4.2.2.4.2 Data Sources

Three data sources were used for the analysis:

- **GDP data**: Data on GDP by MSA was obtained from the U.S. government's Bureau of Economic Analysis (BEA) (Bureau of Economic Analysis, U.S. Department of Commerce 2010). The BEA produces annual estimates of the GDP of each of the 366 U.S. MSAs by computing the sum of the GDP originating in all industries in each MSA.
- **Population data**: Data on the population of each MSA was gathered from the U.S. Census Bureau (U.S. Census Bureau 2010b). The annual MSA population is estimated by the Census Bureau based on the Census 2000 combined with a number of more recent data sources. The Census Bureau points out that because there is a lag in some of the data sources that complement the Census 2000 data, estimates for older vintages tend to be more accurate than those for more recent vintages (U.S. Census Bureau 2008).
- **Data on destinations and frequencies**: This data was prepared using the T100 database which is compiled from data collected by Office of Airline Information (OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b). The T100 database is a complete census of flights by U.S. and foreign

carriers and provides data on the number of operations and passengers carried between each airport pair.

4.2.2.4.3 Defining Hubs

The definition of domestic hubs in the analysis was based on an initial analysis of the T100 database. The objective was to identify those airports that provide connections to the largest number of other airports. For the 2005-2008 time period, the analysis found the number of domestic airports served nonstop[4] presented in Table 4.6, and identified the average number of other OEP-35 airports served nonstop listed in Table 4.7.

---

[4] Only destinations that were served at least 52 times per year were considered, to ensure that at least weekly service existed.

Table 4.6 - Average number of domestic airports served nonstop at least 52 times annually (source: T100 database)

Airport	Average number of domestic airports served nonstop	Rank
ATL	171	1
ORD	141	2
DFW	138	3
MSP	137	4
DEN	134	5
DTW	128	6
IAH	121	7
LAS	119	8
CVG	119	9
CLT	102	10
SLC	101	11

Table 4.7 - Average number of OEP-35 airports served nonstop at least 52 times annually (source: T100 database)

Airport	Average number of OEP-35 airports served nonstop	Rank
ATL	34	1
DEN	34	1
DFW	34	1
MSP	34	1
CVG	33	5
DTW	33	5
IAH	33	5
LAS	33	5
LAX	33	5
ORD	33	5
PHX	33	5

The first four airports in Table 4.7 were connected to all other OEP-35 airports in each of the years from 2005 to 2008. In addition, these airports all rank among the top five airports in terms of the overall number of domestic destinations served, as shown in Table 4.6. The remaining top-five airport from Table 4.6 is ORD which, although it lacks service to one of the OEP-35 airports, ranks as the second most connected airport to other domestic airports. Based on this data, the list of hubs for this analysis is: ATL, ORD, DFW, MSP, and DEN. The impact of this definition is tested as part of the sensitivity analysis discussed in section 4.2.2.6.

4.2.2.4.4 Preparing Benchmark Data

Each of the data sources required some pre-processing for use in the benchmark analysis. This section describes that pre-processing.

Both the GDP and the population data was reported separately for each MSA. Because of the merging of some areas as described in section 4.2.2.4.1, their GDP and population data were summed to provide totals for the entire metropolitan areas.

The data on the number of non-hub destinations served nonstop was computed from data using these conditions and assumptions:

- Departures were considered from the metro area as a whole rather than from individual airports. For instance, if both EWR and LGA airports in the New York region had nonstop service to MSP, this would only be counted as one nonstop destination for the New York metropolitan area.

- At least 52 flights during the year were required in order for a city pair to be considered to have nonstop service.

The data on the daily frequency of service to hubs was prepared using these conditions and assumptions:

- Just as for the number of non-hub destinations served, departures were considered from the metro area as a whole rather than from individual airports. However, in the example with EWR and LGA above, if each airport had service four times daily, the New York region would be counted as having a frequency of eight.
- For those airports that were hubs, only service to the four other hubs could be counted while for non-hub airports, service to the five hubs was counted. To adjust for this, the hub airports' totals were increased by the average of their service to each of the other four hub airports; in practice this amounted to a multiplication of each hub airport's total by a factor of 1.25.

4.2.2.5 Summary of Input and Output Parameters

This section provides four-year average values for each of the four input and output parameters used in the DEA analysis. The full details of the input and output parameters are provided in Appendix E. Although the analysis was done separately for each of the four years, this overview provides averages for the whole period 2005-2008.

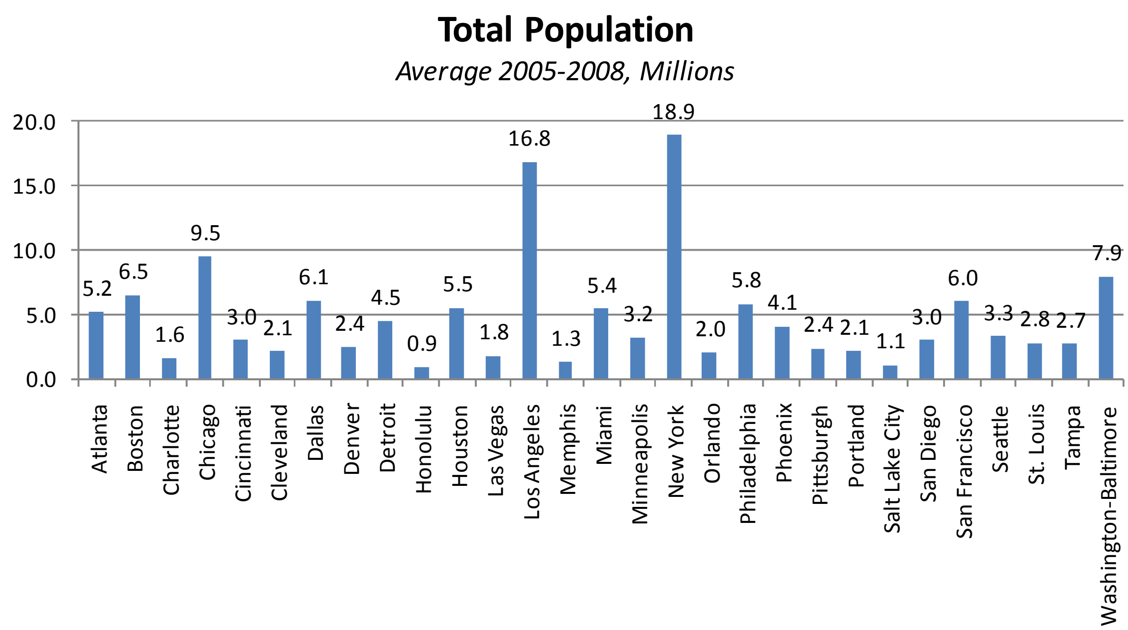

Figure 4.4 - Total population of metropolitan areas in millions, average 2005-2008

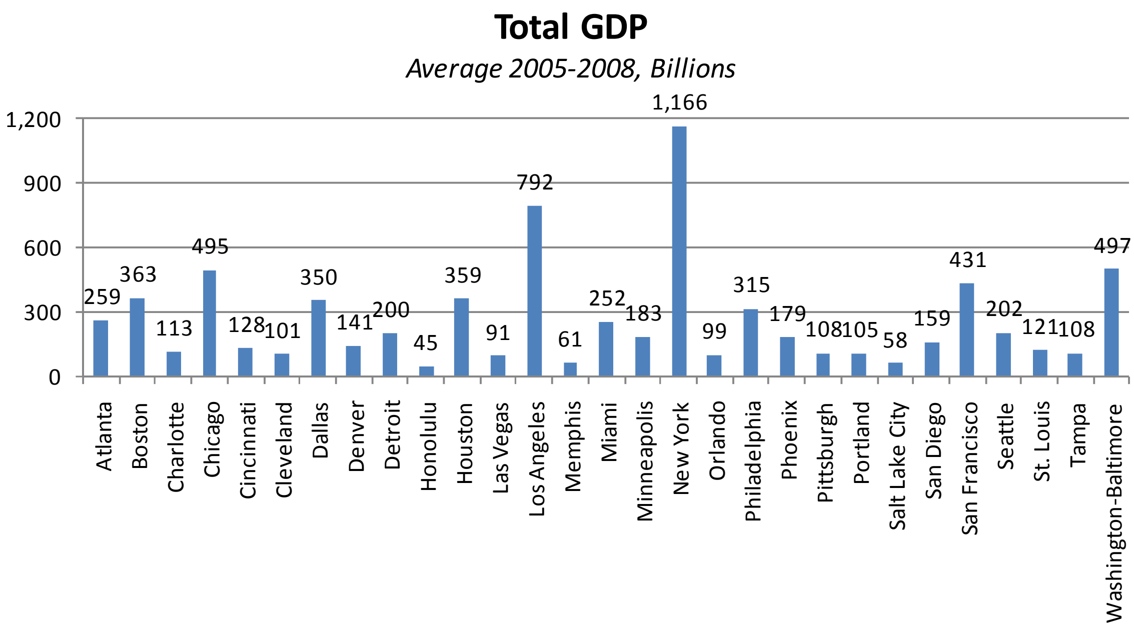

Figure 4.5 - GDP by metropolitan area in billions of US$, average 2005-2008

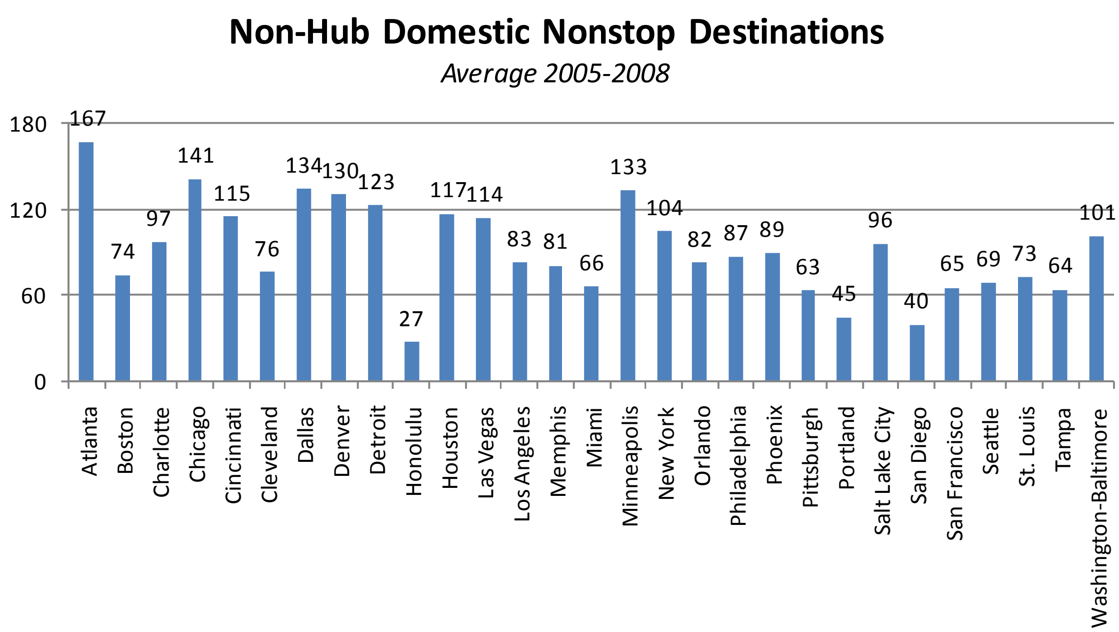

Figure 4.6 –Number of non-hub domestic destinations served nonstop, average 2005-2008

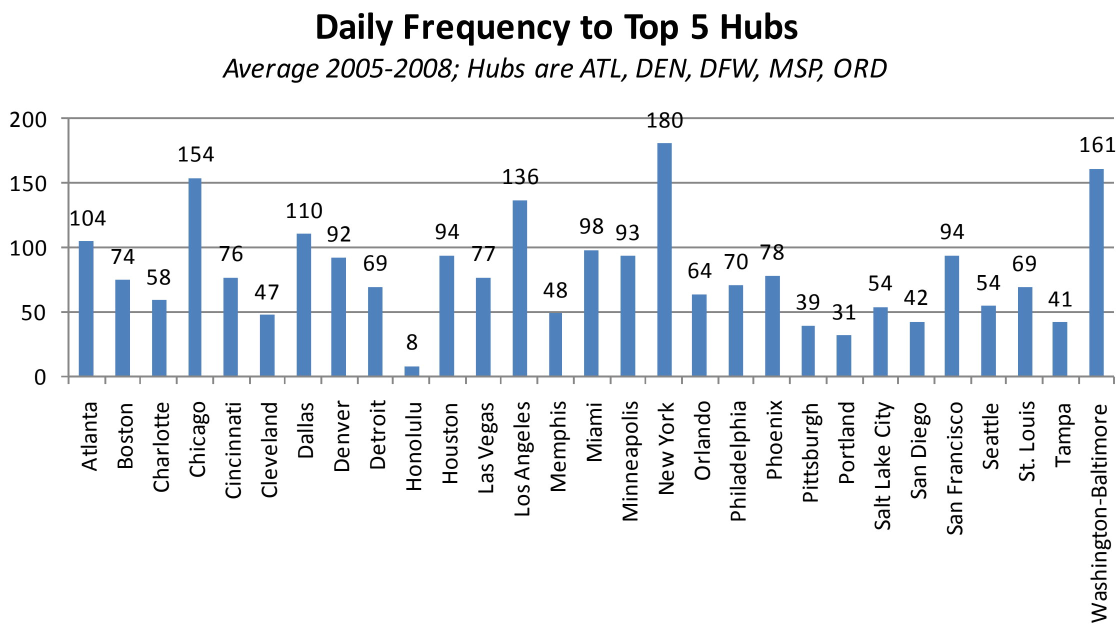

Figure 4.7 - Daily service frequency to top 5 hubs, average 2005-2008

The input data covered those metropolitan areas that have at least one OEP-35 airport. These are the 30 largest metropolitan areas in terms of GDP in the U.S., with the exception of Kansas City, MO, which had on average the country's 28th largest GDP from 2005 to 2008 (Bureau of Economic Analysis, U.S. Department of Commerce 2010) but is not served by an OEP-35 airport and accordingly was left out of the study. Similarly, this represents each of the 30 largest metropolitan areas in terms of population, with the exception of Sacramento, CA, Kansas City, MO, and

San Antonio, TX, which had the 26th, 28th, and 29th largest populations on the average from 2005-2008 (U.S. Census Bureau 2010b).

4.2.2.6 Sensitivity Analysis

The purpose of the sensitivity analysis is to understand the degree to which the findings stand up to any potential changes in the input and output data or the underlying model assumptions of the study.

The choice of DEA model has been shown to have a potentially radical impact on the results of airport performance studies (Schaar & Sherry 2008). Some studies have attempted to address that by using a variety of different models (Sarkis 2000), but this can lead to contradictory and inconclusive results. This paper instead used the framework and heuristics from section 0 to guide model selection. Any variations of the results based on using another DEA model would not be relevant since such a model would be selected without a rationale for its applicability. As a result, no sensitivity analysis using a different DEA model was conducted.

However, in the study of DEA models which use minimum weights, a large body of work exists (e.g. (Mehrabian et al. 2000) and (Allen et al. 1997)) but no conclusive determination of a standard approach to the choice of minimum weights exists. To address this lack of standardization, the sensitivity analysis in this study includes tests of varying these minimum weights.

Regarding the input data on GDP and regional population, no assumptions were made since both of these categories of data were based on government standard definitions. No sensitivity analysis of variations in GDP and population data was conducted. It should however be noted that the analysis results are sensitive to the accuracy of the MSA definitions in terms of their ability to capture a region's entire population and GDP. If an MSA is drawn too wide around a region, it will encompass a larger population and GDP than is actually served by the airport, thereby adversely impacting the region's results in this study. Conversely, if an MSA is drawn too narrowly around a region, the region's results will be over-inflated. Gauging the impact of any MSA boundary errors is not possible in this study.

The data on output parameters regarding the number of non-hub destinations served nonstop and the frequency of service to the top 5 hubs was based not on sampling data but rather on full census data. This means that no sensitivity analysis is necessary to test the impact of sampling errors. However, the data on both of these performance parameters is dependent on the definition of hubs. To test the robustness of the findings with respect to the definition of hubs, the sensitivity analysis included tests of using the top 3, 4, 6, and 7 hubs based on the total number of domestic destinations served nonstop (the list of these airports can be found in Table 4.6).

The results of the sensitivity analysis tests are presented in section 4.2.3.5.

4.2.3 Results

This section presents the resulting scores for the level of air service and discusses the implication of these results. It presents the findings from the sensitivity analysis and discusses some limitations of the results. The section also includes a study of the impact of the level of air service on airline yields.

4.2.3.1 Level of Air Service

The average of the results of the analysis for 2005-2008 is presented in Figure 4.8, where lower scores indicate better levels of service. A k-means cluster analysis was performed on the benchmark results to create the four groups of airports depicted in the figure. The full details of the results are provided in Appendix E. The results are also plotted on a map of the United States in Figure 4.9.

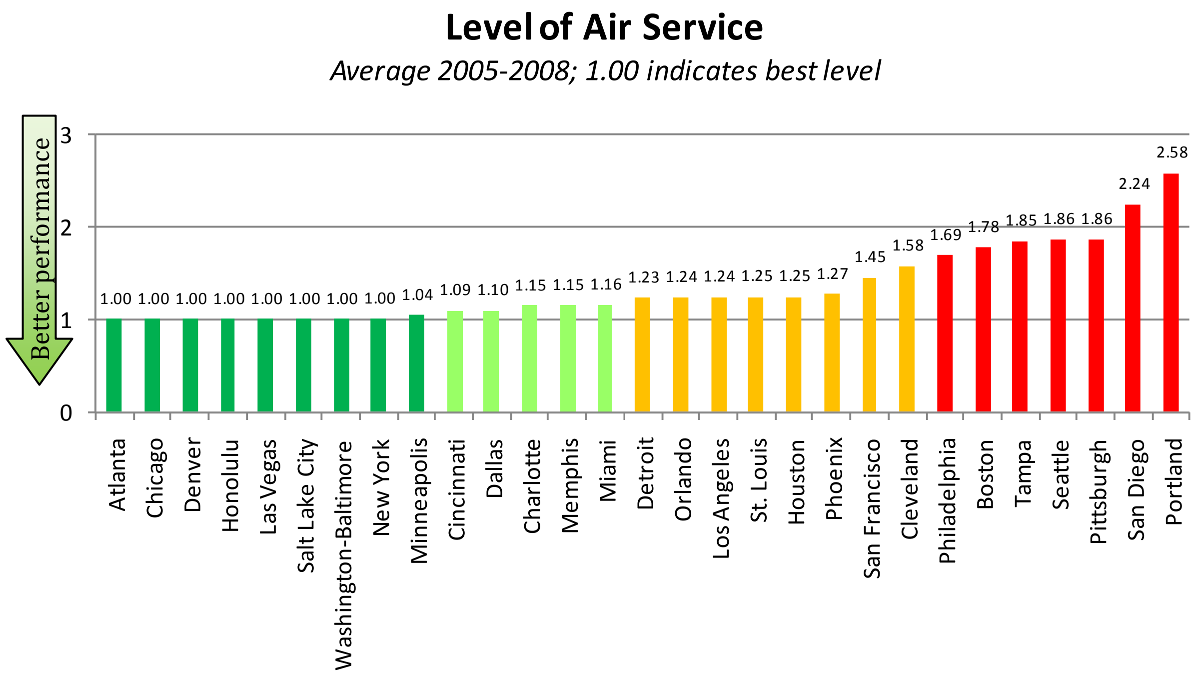

Figure 4.8 – Average levels of Air Service 2005-2008. 1.00 indicates the best level, and high values indicate poor service. Bar coloring is based on a k-means cluster analysis of benchmark scores.

The results show the highest levels of service for Atlanta, Chicago, Denver, Honolulu, Las Vegas, Salt Lake City, and Washington-Baltimore[5]. In contrast, the

---

[5] Note that although New York is listed as 1.00, it is in fact not fully efficient in 2005 but due to rounding error its average appears efficient.

lowest levels of service exist for Portland, San Diego, Pittsburgh, Seattle, and Tampa, with the first two standing out as having lower levels of service.

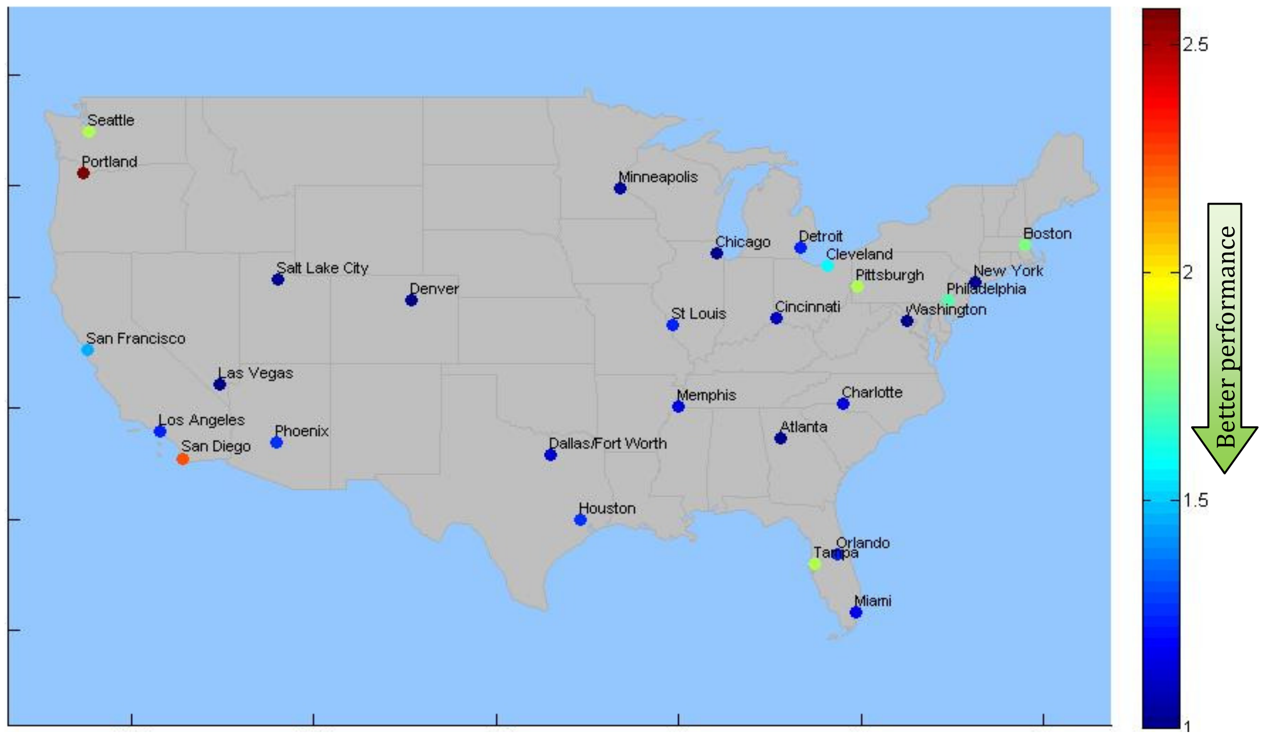

Figure 4.9 - Visualization of levels of air service (Honolulu omitted), 2005-2008 average. 1.00 (dark blue) indicates the best level of air service and high values indicate poor levels of service

4.2.3.2 Gaps for Underserved Metropolitan Areas

The underserved metropolitan areas are defined as those with service levels greater than 1.00, and are considered inefficient in the DEA analysis. The DEA algorithm provides targets which DMUs should hit in order to move from inefficiency to efficiency. The targets are computed by multiplying each output by the DMU's efficiency score from the DEA analysis. These points are the closest projections on the convex hull represented by the efficient frontier.

These projections can provide improvement goals for managers at inefficient airports. When the original parameter values are subtracted from these targets, the gap that must be closed is obtained. Those gaps are presented in Table 4.8. The metropolitan areas in Table 4.8 that have blank values for the gaps for both the number of non-hub nonstops and the number of departures to top hubs are fully efficient in that year.

The inefficient DMUs which have a nonzero slack on one of the output parameters have the shortest distance to the efficient frontier by maximizing output only on the other parameters with a zero slack, irrespective of what is done for the parameter with slack. As a result, the gap for those DMUs to the goal on the frontier is described Table 4.8 only in terms of the parameter with a zero slack, with the other parameter being left blank.

Table 4.8 – Distance to the air service frontier. These are gaps in the level of service to be closed for achieving air service level of 1.00. The gaps are the shortest distance to the frontier.

	Distance to Frontier							
	2005		2006		2007		2008	
	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs
Atlanta								
Boston	56	62	56	60	56	55	61	55
Charlotte	6	4	17	10	22	13	11	6
Chicago								
Cincinnati			11	7	17	12	13	9
Cleveland	37	26	46	30	53	33	37	19
Dallas	12	10	14	12	16	13	12	9
Denver								
Detroit	23	12	32	19	33	19	26	15
Honolulu								
Houston	28	23	30	24	28	23	29	23
Las Vegas								
Los Angeles	22	40	23	38	17	27	19	29

	Distance to Frontier							
	2005		2006		2007		2008	
	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs	Non-hub non-stops	Depts. to top hubs
Memphis	8	5	14	9	12	8	15	9
Miami	14	21	13	20	8	12	6	8
Minneapolis			4	3	11	7	8	5
New York	2	3						
Orlando	21	16	28	21	18	14	12	10
Philadelphia	58	50	63	51	63	50	57	44
Phoenix	31	26	20	19	22	19	23	20
Pittsburgh	38	22	51	29	64	38	51	42
Portland	63	48	68	52	78	52	72	46
Salt Lake City								
San Diego	45	56	43	55	55	53	51	47
San Francisco	33	52	29	43	24	35	32	41
Seattle	66	51	59	47	59	46	53	41
St. Louis	19	17	23	21	17	17	12	13
Tampa	47	29	58	37	57	40	53	34
Washington-Baltimore								

4.2.3.3 Discussion of Results

The results initially show a relatively tight distribution of the levels of service for many airports ranging from 1.00 up to Phoenix at 1.27, where a more severe deterioration occurs, beginning with San Francisco. San Diego and Portland stand out in particular as having worse service than any other metropolitan area. Some factors impacting these results, such as geography, are not controllable, while other factors may be within the scope of influence of airport management and local government.

This section discusses these factors which impact the outcomes of the benchmark. The average levels of air service, GDP per capita, and average gaps are summarized in Table 4.9 along with a brief discussion about the performance of individual metropolitan areas. The remainder of the section discusses the possible causes for high and low levels of air service.

Table 4.9 - Summary of study results, 2005-2008, in order of GDP per capita. Areas with level of air service scores above 1.3 are highlighted in yellow as those areas have poor levels of air service.

Metro Area	GDP/ Capita (Avg.)	Level of Air Service (Avg.)	Distance to Frontier		Comments
			Gap for Destin-ations (Avg.)	Gap for Frequ-ency (Avg.)	
San Francisco	$72,013	1.45	30	43	Somewhat poor air service.
Charlotte	$69,806	1.15	14	8	
Houston	$64,873	1.25	29	23	
Washington-Baltimore	$62,526	1.00	0	0	Full air service
New York	$61,692	1.00	0	1	Nearly full air service (rounding error)
Seattle	$61,652	1.86	59	46	Poor air service. Located in the far Northwest where no metropolitan area has high levels of air service.
Denver	$58,004	1.00	0	0	Full air service
Dallas	$57,555	1.10	13	11	
Minneapolis	$57,473	1.04	6	4	

Metro Area	GDP/ Capita (Avg.)	Level of Air Service (Avg.)	Distance to Frontier		Comments
			Gap for Destin-ations (Avg.)	Gap for Frequ-ency (Avg.)	
Boston	$55,893	1.78	57	58	Poor air service in spite of including BOS, PVD, and MHT in this metropolitan area. One factor is that PVD is heavily dominated by Southwest Airlines (American University School of Communication 2010) which results in limited service to the top hubs.
Philadelphia	$54,163	1.69	60	49	Poor air service.
San Diego	$53,630	2.24	49	53	Poor air service, and only a single runway.
Salt Lake City	$53,216	1.00	0	0	Full air service
Chicago	$52,196	1.00	0	0	Full air service
Las Vegas	$50,998	1.00	0	0	Full air service
Atlanta	$50,016	1.00	0	0	Full air service
Honolulu	$49,869	1.00	0	0	Full air service
Orlando	$49,146	1.24	20	15	In spite of extensive holiday traffic, Orlando is not at full air service.
Portland	$48,888	2.58	70	49	Poor air service. Low yields may contribute (see Figure 4.13).

Metro Area	GDP/ Capita (Avg.)	Level of Air Service (Avg.)	Distance to Frontier		Comments
			Gap for Destin-ations (Avg.)	Gap for Frequ-ency (Avg.)	
Cleveland	$48,269	1.58	43	27	Somewhat poor air service. Reduction of hubbing by Continental may contribute (Rollenhagen 2003).
Memphis	$48,056	1.15	12	7	
Los Angeles	$47,074	1.24	20	33	
Miami	$46,632	1.16	10	15	
Pittsburgh	$45,638	1.86	51	33	Poor air service, in large part due to US Airways hub elimination (Grossman 2007). Service deteriorated each year from 2005 to 2008.
Detroit	$44,756	1.23	29	16	
Phoenix	$43,828	1.27	24	21	
St. Louis	$43,217	1.25	18	17	
Cincinnati	$43,040	1.09	10	7	
Tampa	$39,932	1.85	54	35	Poor air service. The city's relative proximity to Orlando could contribute, but that impact should be limited since Tampa city center is 86 miles from MCO.

4.2.3.3.1 Impact of Geography

Although many of the less well served metropolitan areas are located in one of the four “corners” of the continental United States as shown in Figure 4.9, many of these less well served metropolitan areas exist in the vicinity of other metropolitan areas with high levels of service. This suggests that some areas’ lower levels of service may stem less from their geographic distance from the center of the country and more from their proximity to another well-served metropolitan area.

For example, Tampa exhibits low levels of air service and is located in the southeast corner of the United States, but neighboring Orlando exhibits high levels of air service. This suggests that Tampa’s low level of air service may be traced more to its proximity to Orlando than to its southeasterly location.

Seattle and Portland are exceptions to this, since they both exhibit low levels of service and are not in the proximity of a well-served area.

4.2.3.3.2 Impact of Capacity Limitations

A lack of infrastructure capacity in the form of runways, terminals, or other facilities at an airport may limit the ability of airlines to add service even though demand exists.

A proxy for capacity limits is the level of delays at an airport; heavy delays suggest that the airport infrastructure has difficulty accommodating the level of demand at the airport. Data on on-time arrivals at the airports in the study was obtained from the BTS airline on-time database (Bureau of Transportation Statistics 2010a). A weighted average was computed for each area, and the average included all airports located in the area, weighted by the number of operations at each airport. Figure 4.10 shows the percentage of on-time arrivals in relation to the benchmark results for the level of air service.

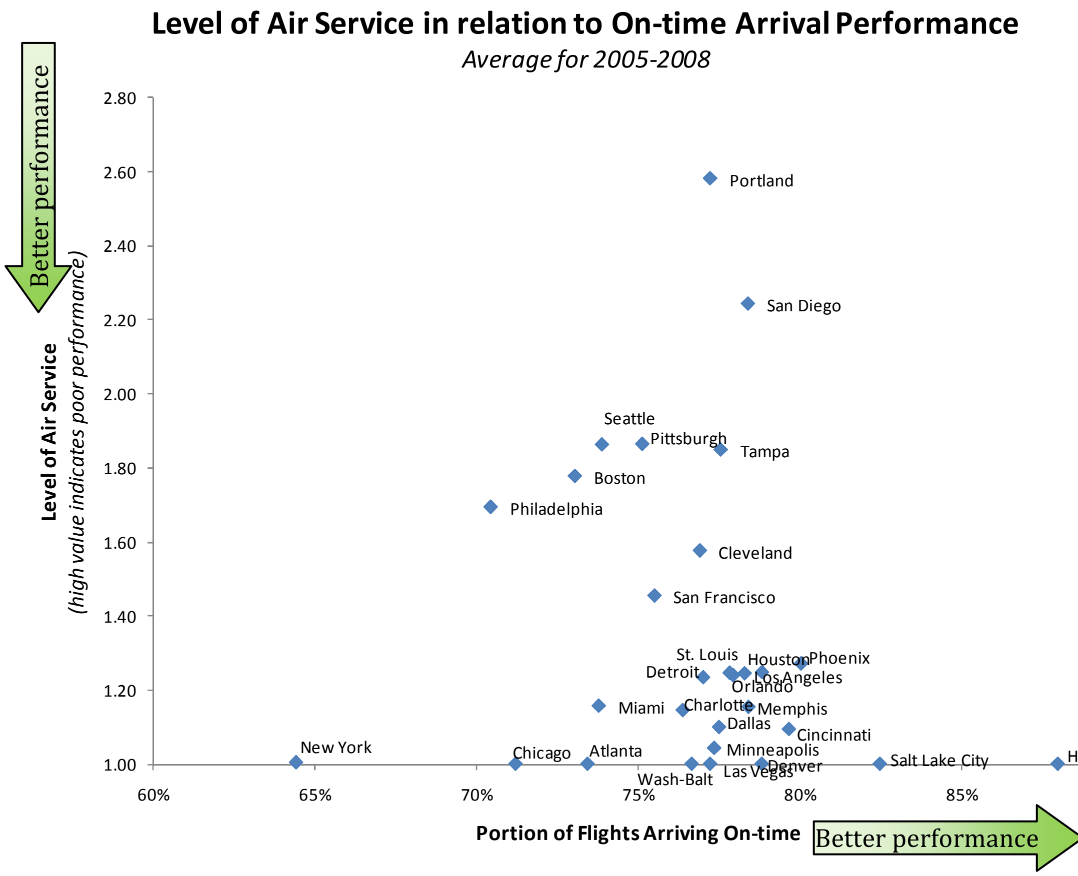

Figure 4.10 - Percentage of airline on-time arrivals in U.S. metropolitan areas; average for 2005-2008. On-time arrivals for regions with multiple airports are weighted by the volume of operations at each airport in the region.

This data suggests that a contributing cause of the low levels of air service in areas such as Philadelphia, which has the second-worst arrival delays, may be capacity limitations. Other areas such as New York and Chicago are currently well-served in terms of the level of air service, but because of capacity limitations, they may find that the future level of air service cannot grow at the same level as their population and regional economies, resulting in a proportionately reduced level of air service.

To address capacity limitations, airports and the FAA can undertake projects to improve the capacity. These capacity improvements are summarized in (Federal Aviation Administration 2007, p. 27 of app. C in source document) and Table 4.10 shows the improvements which will be applied prior to 2015 and 2025, respectively. Note that the table omits three improvements[6] which will be made to all OEP-35 airports, as those improvements do not provide any distinguishing approaches to each airport.

---

[6] These three improvements are: 1) Reduced Separation Standards (use visual separation in MMC and use 2/3/4/5 NM in IMC); 2) Improved threshold delivery accuracy; 3) 1.5 NM Departure/Arrival separation (IMC) (spacing < 2500 ft or same runway)

Table 4.10 - Planned airport capacity improvements (Federal Aviation Administration 2007). For greyed-out airports no data was available. Capacity improvements that apply to all OEP-35 airports have been omitted.

Metro Area	Airport	New/extended runways (2006 or later)	Independent parallel approaches (IMC) -- spacing 2500-4299 ft	Triple indep. parallel approaches (IMC)	Mixed triple independent/dependent parallel approaches (IMC)	Paired approaches, e.g. SOIA -- MMC (spacing 700-2499 ft)	-- IMC (spacing 1200-2499 ft)	Dependent Approaches -- MMC/IMC (700-2500 ft spacing) -- 1.5 NM diagonal behind Small, Large -- wake vortex sep behind B757/Heavy	LAHSO (all weather) if >7000 ft to intersection	Simultaneous Converging Approaches (IMC)	Standard Departure/Departure separations (no departure constraints)	Independent parallel departures (IMC) -- no wake vortex separation behind Small/Large (700-2500 ft spacing)
Atlanta	ATL					< 2025						
Boston	BOS	< 2015				< 2015	< 2025				< 2025	
	MHT											
	PVD											
Charlotte	CLT	< 2025			< 2025							
Chicago	MDW											
	ORD	< 2015		< 2015								
Cincinnati	CVG											
	DAY											
Cleveland	CLE											
Dallas	DAL											
	DFW	< 2025										
Denver	DEN	< 2025										
Detroit	DTW			< 2025								
Honolulu	HNL								< 2025			
Houston	HOU	< 2025										
	IAH	< 2025										
Las Vegas	LAS					< 2025			< 2025	< 2025	< 2025	< 2025
Los Angeles	LAX					< 2025		< 2025				
	ONT											
	BUR											
	SNA											
	LGB											
Memphis	MEM											
Miami	FLL	< 2015	< 2025								< 2025	

Metro Area	Airport	New/extended runways (2006 or later)	Independent parallel approaches (IMC) -- spacing 2500-4299 ft	Triple indep. parallel approaches (IMC)	Mixed triple independent/dependent parallel approaches (IMC)	Paired approaches, e.g. SOIA -- MMC (spacing 700-2499 ft)	-- IMC (spacing 1200-2499 ft)	Dependent Approaches -- MMC/IMC (700-2500 ft spacing) -- 1.5 NM diagonal behind Small, Large -- wake vortex sep behind B757/Heavy	LAHSO (all weather) if >7000 ft to intersection	Simultaneous Converging Approaches (IMC)	Standard Departure/Departure separations (no departure constraints)	Independent parallel departures (IMC) -- no wake vortex separation behind Small/Large (700-2500 ft spacing)
	MIA					< 2025		< 2025	< 2025			< 2025
	PBI	< 2015										
Minneapolis	MSP									< 2025		
New York	JFK										< 2015	
	LGA											
	EWR					< 2015		< 2025				< 2025
	ISP											
Orlando	MCO			< 2015				< 2025				< 2025
Philadelphia	PHL	< 2015										
Phoenix	PHX		< 2025									
Pittsburgh	PIT											
Portland	PDX		< 2025			< 2015					< 2025	
Salt Lake City	SLC			< 2025							< 2025	
San Diego	SAN										< 2025	
San Francisco	SFO							< 2025			< 2025	< 2025
	SJC											
	OAK					< 2025		< 2025		< 2025		< 2025
Seattle	SEA	< 2015	< 2025			< 2015						
St. Louis	STL							< 2015				< 2015
Tampa	TPA	< 2025						< 2025				< 2025
Washington-Baltimore	DCA											
	IAD	< 2015 < 2025		< 2015								
	BWI	< 2025										

The table shows improved capacity planned at PHL in the form of new/extended runways but no new/extended runways at any of the airports in Los

Angeles, New York or San Francisco. It also shows no planned improvements at LGA.

4.2.3.3.3 Impact of a Lack of Hub Service

An airport's status as a hub for a carrier brings connecting passenger traffic, allowing the air carrier to provide higher frequency service and to serve more destinations than would have been possible if the airport served primarily O&D traffic (Belobaba et al. 2009, p. 163).

Pittsburgh's low level of air service is in part the result of its lack of hub status for any airline since US Airways consolidated its hubs to Philadelphia and Charlotte (Grossman 2007). Similar conditions may exist in Cleveland (Rollenhagen 2003) which also reports a relatively low level of air service.

4.2.3.3.4 Impact of Local Industry Base

The needs for air transportation may vary by industry. For instance, in a comparison of two areas with the same GDP and population, it may be that one has better conditions for higher levels of air service than the other as a result of differences in industry makeup.

In a 2000 study, it was found that regions with a stronger focus on high-tech industries were more likely to have airline hub service (Button & Stough 2000, pp. 231-264). The industries that were defined as part of the high-tech industry included mining; construction; chemicals; fabricated metal products; electronic equipment; communications; financial services; real estate; business, engineering and management services; and many others. The definition of high-tech industries used in that study was derived from a 1986 characterization of high-tech industries (Rees 1986, pp. 76-92), which determined that the high-tech industry was made up of 88 categories of economic activity from the Standard Industrial Classification (SIC) system. The SIC system was replaced in 1997 by the North American Industry Classification System (NAICS) (U.S. Census Bureau 2010a), and data is not publicly available at the detailed level of the 88 original categories for the 2005-2008 time period.

To provide a high level assessment of the high-tech industries' role in the level of air service of metropolitan areas, the original 88 categories were consolidated into five high-level categories from the NAICS: Mining; utilities; manufacturing; information; and professional, scientific, and technical services. The categories are described in Table 4.11.

Table 4.11 - Definition of NAICS categories included in high-tech industry (U.S. Census Bureau 2002)

Industry	Definition
Mining	Mining is defined as establishments that extract mineral solids, liquid minerals, and gases. The industry includes both mine operations and mining support activities.
Utilities	Utilities include establishments that provision electric power and natural gas, supply water and steam, and remove sewage.
Manufacturing	Manufacturing encompasses establishments that perform mechanical, physical, or chemical transformation of materials, substances, or components into new products. The manufacturing category encompasses many types of manufacturing, such as food, paper, petroleum, and plastics manufacturing.

Industry	Definition
Information	Information establishments are engaged in "(a) producing and distributing information and cultural products, (b) providing the means to transmit or distribute these products as well as data or communications, and (c) processing data" (U.S. Census Bureau 2002).
Professional, scientific and technical services	This category "comprises establishments that specialize in performing professional, scientific, and technical activities for others. These activities require a high degree of expertise and training" (U.S. Census Bureau 2002).

Figure 4.11 presents the data retrieved using this definition, and Figure 4.12 shows a scatter plot of the level of air service in relation to the high-tech industry contribution as a portion of the total regional economy. Since a normality assumption could not be made about the benchmark results data, a Kruskal-Wallis test was conducted to compare the benchmark results of the half of metropolitan areas with a low portion of high-tech industry to the areas which have a high portion of high-tech industry. The results of the test are presented in Table 4.12.

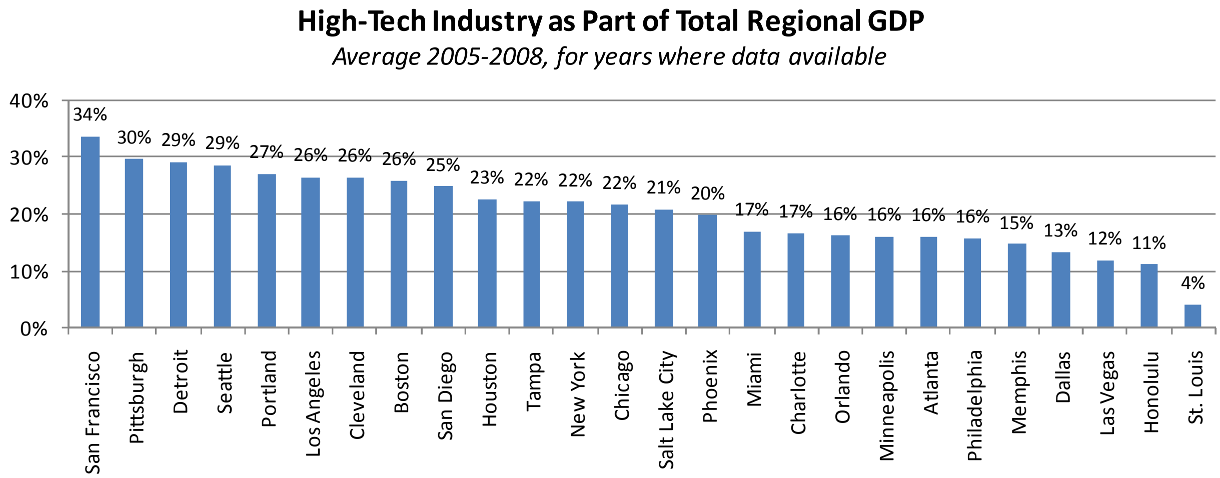

Figure 4.11 – High-tech industry as a percentage of total regional GDP. Cincinnati, Denver, and Washington-Baltimore omitted due to lack of data. High-tech industry defined as NAICS categories of mining, utilities, manufacturing, information, and professional, scientific and technical services. Average based only on those years where data in the largest number of categories was available for each city; some categories are marked in the data source as "not shown in order to avoid the disclosure of confidential information". (Bureau of Economic Analysis, U.S. Department of Commerce 2010)

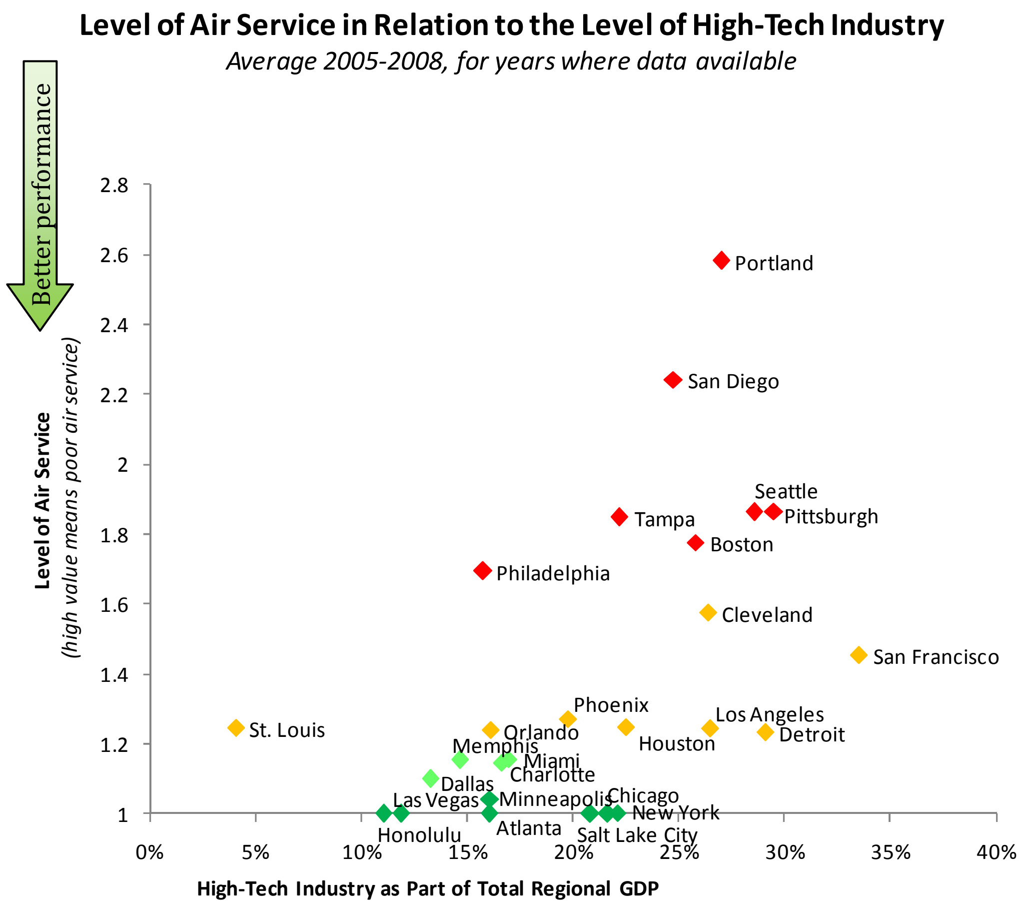

Figure 4.12 - Level of air service as a function of high-tech industry as a percentage of total regional GDP (methodology description in caption for Figure 4.11). Marker coloring is based on a k-means cluster analysis of benchmark scores.

Table 4.12 - Results of Kruskal-Wallis test of the benchmark results based on grouping

Years	Mean rank		Chi-square	Asymptotic significance
	High portion high-tech	Low portion high-tech		
2005-2008	17.46	8.77	7.579	0.00591

Instead, the results indicate that the areas with poor levels of air service exhibit high portions of high-tech industry; Seattle and Portland are such examples. This is different from the findings of (Button & Stough 2000). Two key differences exist between the study of (Button & Stough 2000) and the present analysis: First, the definition of high-tech industry differs between the two studies as a result of the change from the SIC to the NAICS system. Second, approximately ten years separate the two studies, suggesting that the characteristics of the different industries may have changed during that time. Further analysis of more detailed GDP breakout data is necessary to find if other industries have an impact on the level of air service.

4.2.3.3.5 Impact of Airline Yield

Airlines are private enterprises which seek to make a profit by supplying services to meet demand. Where profits are high, the incentive exists for air carriers to add more service, while locations where profits are low provide less incentive for increased levels of service. The level of air service in a metropolitan area may be dependent on the level of airline yields for services to and from that area.

The average yields were computed for the 2005-2008 period for each airport in the analysis. Data on revenues and passenger volumes were derived from the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c). Annual yields, expressed in US$ per Revenue Passenger Mile (RPM) for each O&D pair is computed as follows (Belobaba et al. 2009, p. 48):

$$Yield_{O,D} = \frac{Revenue_{O,D}}{Distance_{O,D} * Passengers_{O,D}}$$

The yields for each metropolitan area was determined by computed a passenger-weighted yield from the data from each airport in the area. The resulting data is displayed in Figure 4.13.

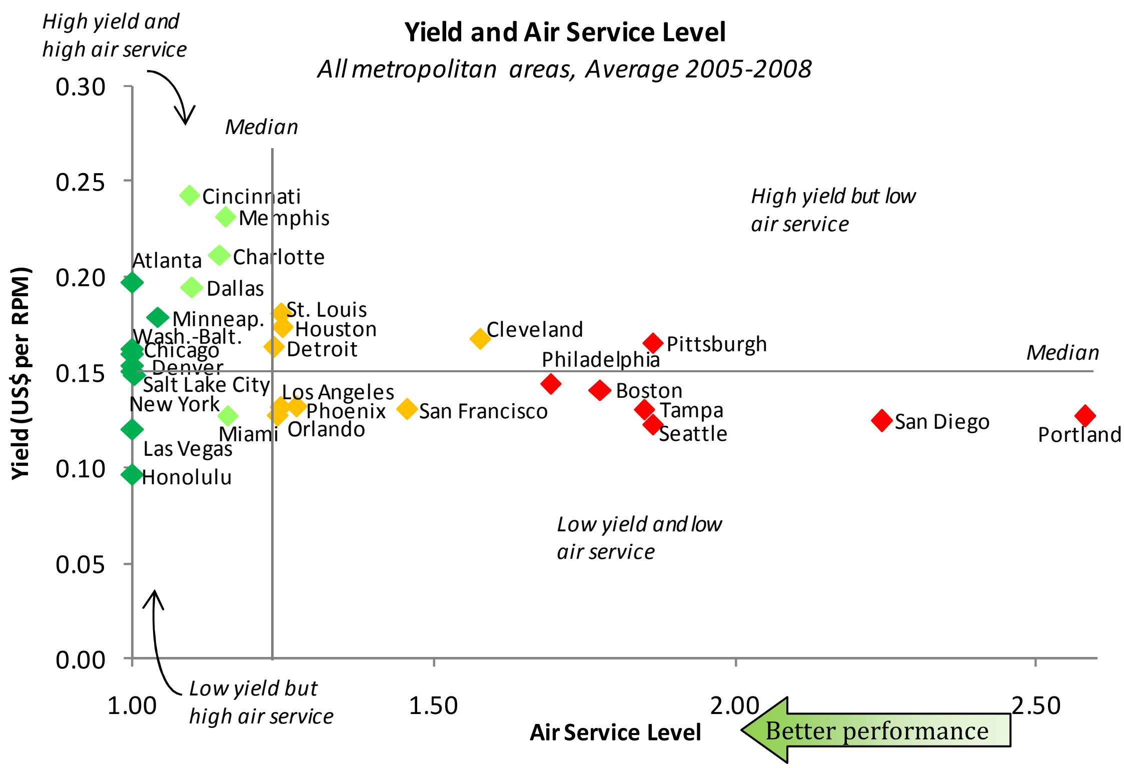

Figure 4.13 – Yield by metropolitan area in relation to air service level. The best air service level is at 1.00 and the worst service level is at 3.00. Marker coloring is based on a k-means cluster analysis of the benchmark results.

The data in Figure 4.13 suggests that low yields may be contributing to low levels of air service in areas such as Portland and San Diego, located in the lower right-hand quadrant in the figure. Low demand or high degrees of competition in these areas contribute toward the low yield, making them less attractive to airlines for adding further services.

The data also suggests that some areas exist where the level of air service is high in conjunction with high yields. This is the top-left quadrant of the figure and includes areas such as Cincinnati and Memphis. The focus for the airport authority and local government in these areas should not be to improve air service but rather to create a more competitive environment where yields are reduced as a result of lower fares for travelers.

Lastly, the areas located in the upper right-hand quadrant of the figure are under-served in terms of air service but report above-median yields. These appear to represent opportunities for air carriers in that there is room for adding new, profitable service. The airport authority and local government in these areas should focus on both adding increased service and creating more competition in order to reduce travel costs for its residents. These cities include St. Louis, Houston, Detroit, Cleveland, and Pittsburgh.

4.2.3.4 Limitations of Results

- Although the extent of their impact is unknown, several factors which may have affected the outcome of the analysis exist:
- The calculation of the level of air service does not factor in the geographic location of the metropolitan area. It is possible that areas located near the center of the continental United States have an inherently greater possibility of achieving high levels of air service.
- The calculation does not take into account the effects of economic geography. It is for instance possible that the industrial base of one metropolitan area is more prone to using air service than that of other areas.
- The calculation does not account for the impact of capacity limitations on gates, runway capacity, etc. It appears that this limits the score for cities like San Diego.
- The calculation does not consider the relatively close proximity of some metropolitan areas to other areas. It is possible that the proximity to another area impacts a region's level of air service.
- International traffic was excluded from the study since 14 airports among the OEP-35 airports represented 70% of all international

passenger enplanements in 2006 (FAA 2008, pp. 23-24). A study that included international traffic would show different results.

- The sensitivity analysis could not address whether the MSA boundaries were drawn too wide or too narrowly around any regions.

4.2.3.5 Results of Sensitivity Analysis

This section presents the results of the sensitivity analyses described in section 4.2.2.6.

4.2.3.5.1 Sensitivity to Weight Boundaries

In the original analysis, the standard weight boundaries ε from the BCC model were used. These are the boundaries on the minimum values on the weights applied to each output in the DEA calculation. In the BCC model these are simply specified as infinitesimal and in the model implementation, they were set at 1.0 * E-6.

In the sensitivity analysis, the weights were varied between the minimum value of 1.0 * E-6 up to the maximum feasible output weight values. The maximum feasible values are the maximum observed values multiplied by 0.5 (as a result of there being two output parameters). The maximum feasible values are those which result in the constraints being binding for one or more DMUs.

The input parameter weights are not varied since any minimum values unfairly penalize the performance of the larger metropolitan areas due to the differences in magnitude of the different areas' values.

In the case where the analysis uses the maximum feasible weights, the DMU(s) with the highest magnitude of outputs are forced to apply exactly those weights, effectively removing the DMU's ability to select its own optimal weights. The higher the boundary on weights, the lower the flexibility for DMU's to determine their own optimal weights.

For the output weights, seven variations on the weight boundaries were tested for each year; the first test i=1 used the standard 1.0 E-6 weights, and in each subsequent test i=2..7 the boundary was proportionally increased such that the test i=7 had the maximum feasible boundaries (for tests i=2..7 the weight boundaries were determined as $\text{boundary}_i = \max(\text{weight}) / 2 * (i - 1) / 6$).

The average scores computed in the sensitivity analysis are presented in Figure 4.14. A comparison of the rankings of each metro area's scores between Test 1, Test 2, and Test 7 is presented in Table 4.13.

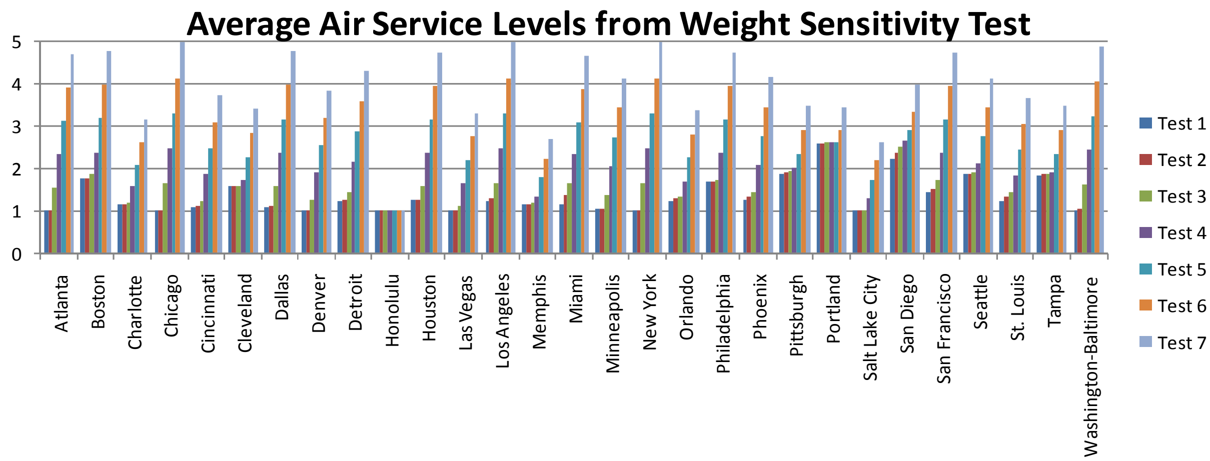

Figure 4.14 - Results from weight boundary sensitivity tests. Test 1 has the least restrictive weight boundaries, and Test 7 has the most restrictive boundaries

Table 4.13 - Rankings from selected sensitivity tests. Test 1 has the least restrictive weight boundaries and Test 7 has the most restrictive boundaries.

	Ranking in Sensitivity Test			
	Test 1	Test 2	...	Test 7
Atlanta	1	1		20
Boston	24	24		25
Charlotte	12	12		4
Chicago	1	1		27
Cincinnati	10	11		12
Cleveland	22	22		7

	Ranking in Sensitivity Test			
	Test 1	Test 2	...	Test 7
Dallas	11	10		24
Denver	1	1		13
Detroit	15	15		18
Honolulu	1	1		1
Houston	19	14		21
Las Vegas	1	1		5
Los Angeles	17	17		27
Memphis	13	13		3
Miami	14	20		19
Minneapolis	9	8		15
New York	8	7		27
Orlando	16	16		6
Philadelphia	23	23		22
Phoenix	20	18		17
Pittsburgh	27	27		10
Portland	29	29		8
Salt Lake City	1	1		2
San Diego	28	28		14

	Ranking in Sensitivity Test			
	Test 1	Test 2	...	Test 7
San Francisco	21	21		23
Seattle	26	26		16
St. Louis	18	19		11
Tampa	25	25		9
Washington-Baltimore	1	9		26

The results show that the rankings in Test 1 and Test 2, which have the lowest boundaries, remain largely the same. Between the two tests, 20 metropolitan areas retain the same ranking, 6 areas shift one or two rankings, and 3 areas shift more than 2 rankings. However, with Test 3, rankings begin to shift more drastically, and by Test 7 only one airport, Honolulu, maintains its original ranking.

This shows that the selection of weight boundaries do matter to the results if they go well above the infinitesimal. However, the BCC model specifies that infinitesimal weight boundaries be used, and the similarity between the results of Test 1 and Test 2 shows that the exact choice of infinitesimal weight boundaries in the model implementation has little impact; the boundaries in Test 2 already far exceed what could be considered reasonable infinitesimal weight boundaries in the model. This indicates that the boundaries of 1.0 * E-6 used in the analysis are acceptable.

4.2.3.5.2 Sensitivity to Hub Definition

In the sensitivity test where the definition of hubs was changed as described in section 4.2.2.6, tests were run for 3, 4, 5, 6, and 7 hubs. The results were then averaged across all cases, and standard deviations for the level of air service were computed. The results of this analysis are shown in Figure 4.15.

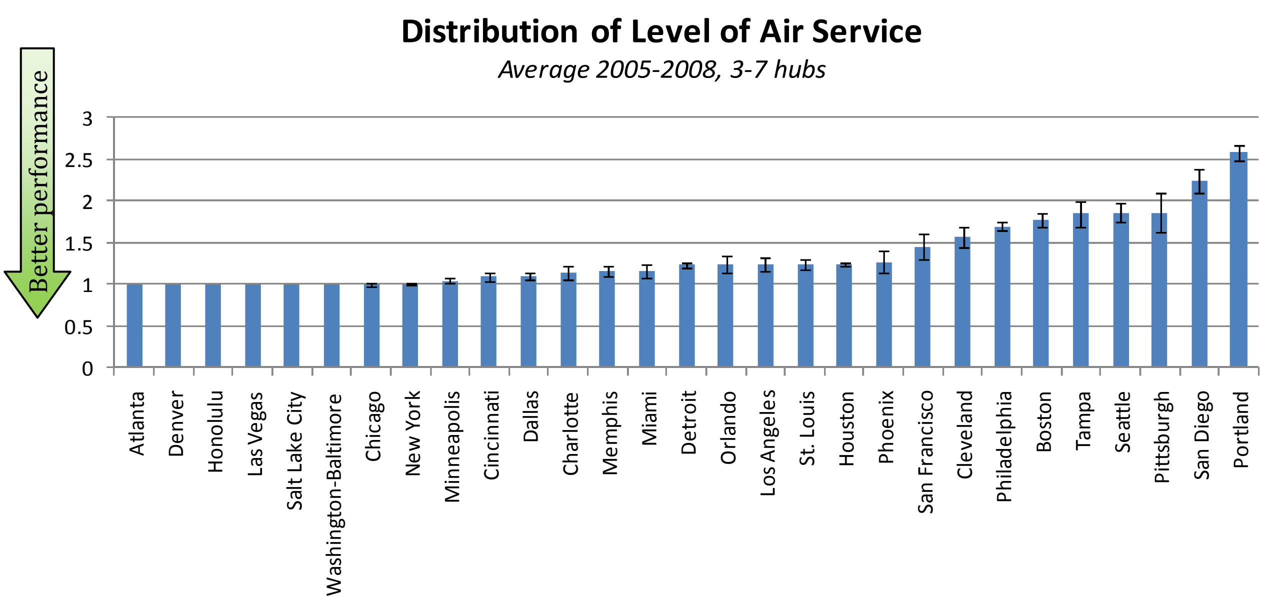

Figure 4.15 - Results from hub definition sensitivity test. The bars show the average score, and the error bars show +- the standard deviation.

The results show a very small standard deviation for the fully efficient metropolitan areas and limited standard deviations elsewhere, indicating a very limited impact on the results from changes in the definition of how hubs are determined.

4.2.4 Conclusions

The analysis defined the level of service in U.S. metropolitan areas as the number of non-hub destinations served and the frequency of service to hubs, and found that the least well-served areas are Portland, OR, San Diego, CA, Pittsburg, PA, and Seattle, WA. The analysis presented the gaps that, if closed, would have resulted in matching the level of air service at the best-served areas.

The results suggest that some areas have a lower opportunity for high levels of air service as a result of their geographic proximity to other areas with high degrees of air service or as a result of their regional industry base generating a low level of demand for air travel. These are factors that are generally uncontrollable by local airport authorities or local and federal government bodies.

In contrast, the results suggest that some areas have a low level of air service as a result of factors which may be controlled or influenced.

Areas such as Philadelphia appear to be underserved as a result of limited infrastructure capacity. This suggests that allocating funding for adding new capacity through for instance a new runway should result in improved service for the local population and economy.

Other areas, such as Pittsburgh and Cleveland report low levels of air service but also record above-median yield levels. This suggests that these areas represent opportunities for added service by air carriers, and that the local airport authorities and government should focus efforts on recruiting new air service in order for the region's population and businesses to be better served.

4.3 Case Study 2: Benchmark of the Level of Capacity Utilization at U.S. Airports

Airport capacity utilization has in the past been defined in terms of the number of aircraft handled by each unit of runway capacity (Federal Aviation Administration 2004). While this measure is an important indicator of how well a scarce resource is being utilized, this case study studies whether that metric is sufficient for addressing the concerns about capacity utilization of several key airport stakeholders. The case study leverages the analysis of airport stakeholders and their goals to establish new measures of how well the airport's capacity is being used to meet the goals of a broader set of stakeholders.

As section 4.2.1 shows, U.S. airports exist to serve the needs of a broad range of stakeholders. The airports provide access to air transportation services to a region's residents, visitors, and businesses. It is the objective of these stakeholders and of several other stakeholders, that the airport capacity be used to maximize the level of air service and that as many passengers as possible be carried through the airport's infrastructure.

This analysis provides a benchmark of the degree to which U.S. airport capacity is being used to meet these stakeholder goals. The key findings include:

- Seven of the 35 airports exhibit full efficiency in their level of capacity utilization. Some of these airports currently operate at high levels of delay, suggesting that obstacles exist for continued growth in capacity utilization

- Although increasing the average size of aircraft used at airports with currently high capacity utilization could help further increase the level of capacity utilization, limited demand to markets currently served curbs the potential impact of increased aircraft size unless air carriers also switch to serving higher-demand markets.
- Six of the 35 airports exhibit poor levels of air service. These airports are PDX, PIT, TPA, HNL, MCO, and MDW. Some of these airports exist in currently under-served markets, suggesting that capacity utilization could be improved, while others are in already well-served markets.

Note that in this case study the analysis is based on individual airports, in contrast to case study 1 that was presented in section 4.2 which was based on metropolitan areas that often included more than one airport.

This case study is organized as follows: Section 4.2.1 introduces the model of airport stakeholders and their goals. Section 4.2.2 describes the study methodology. Section 4.2.3 presents the study results. Section 4.2.4 discusses the conclusions.

4.3.1 The Airport's Stakeholders and Their Goals

The analysis of airport stakeholders in section 2.1.3 found that the stakeholders' goals for the airport were based in part on factors wholly within the control of airport management (the "airport organizational boundary"), but also on factors that were only partly within the control of management, or entirely outside management's control.

The goal of "maximizing the number of destinations served and frequency of those services" emerged from the analysis as common to stakeholder groups such as local businesses, residents, the local government, and the airport organization itself. It is an example of a goal that is not fully within the control of airport management since airlines determine where to add or reduce service.

Several stakeholders also have an interest in maximizing the number of passengers carried through the airport. For instance, local businesses dependent on tourism benefit from maximizing passenger throughput. Concessionaires at the airport can see increased revenues through larger volumes of passengers. Airport management have an incentive in ensuring maximum passenger throughput since they bring increased revenues to the airport both in the form of Passenger Facility

Charges (PFCs) and concessions spending which results in increased non-aeronautical revenues for the airport.

The goals reflects a “symbiotic” relationship between a region’s economy and the local air service, where air service stimulates economic growth (Button & Stough 2000) and growth in a region’s economy drives increased demand for air travel.

The goals may be in conflict with each other. Airlines can ensure that the number of destinations served is maximized by using aircraft that are sized appropriately for the levels of demand for each market. Although the number of passengers carried might be maximized by flying larger aircraft to certain markets where demand is high, airlines may have an incentive to fly smaller aircraft to a larger number of destinations thanks to higher levels of yield to these small markets. At capacity-constrained airports, a conflict exists then exists between achieving the objective of maximizing the number of destinations served and maximizing the number of passengers carried.

The stakeholders who are concerned with these potentially conflicting goals have a need for evaluating the degree to which it is being achieved in U.S. metropolitan areas. For instance, local governments and airport authorities must understand if their region is currently well served by airlines or if added effort is necessary to attract additional air service. If a shortfall exists in the degree to which the goal is being met, they must gain insight about what is causing the performance gap. Conversely, a region's residents and business community must understand if their needs are being met by the airport(s) in their region, or if they should demand more from their local government and airport authority in terms of attracting new air service to their community.

A comparative benchmark is a means to evaluate these goals. The benchmark allows for a normalized comparison across major U.S. metropolitan areas and gives stakeholders an understanding of airports which are not effectively meeting goals and can also provide insight into the causes of any performance gaps. The evaluation of stakeholder goals indicate that what should be benchmarked is the level to which airport capacity is used to provide service to a large number of destinations; to provide high frequency of service; and to transport a large number of passengers.

4.3.2 Methodology

This section discusses the study methodology. It provides the motivation for the selection of performance parameters and discusses the choice of model for benchmarking. It also describes the data sources and pre-processing as well as the method used for computing benchmark scores. Finally, it presents the method for sensitivity analysis of the results.

4.3.2.1 Scope of Analysis

As described in section 1.1, the scope of this analysis is limited to the U.S. OEP-35 airports only.

To ensure accurate comparisons, the scope of the study is limited to domestic U.S. air service only. International traffic was excluded from the study since 14 airports among the OEP-35 airports represented 70% of all international passenger enplanements in 2006 (FAA 2008, pp. 23-24). The details of how this limited scope was implemented are discussed in subsequent sections relating to data pre-processing.

4.3.2.2 Selection of Model Parameters

The analysis of stakeholder goals in section 4.2.1 described their goals for utilizing airport capacity. This section reviews the details of these goals and translates them into specific performance parameters.

The stakeholder goals were combined into the conceptual ratio of (number of destinations served; high frequency of service; number of passengers) : (airport capacity). The first three parameters can be thought of as outputs from the process while the last parameter, airport capacity, is the resource input that makes the outputs possible. As airport capacity increases, assuming the presence of sufficient demand, the expectation is that the “production” of outputs would increase in the form of destinations served, frequency of service, and/or volume of passengers served.

4.3.2.2.1 Measuring the Level of Air Service

The goal includes maximizing both the number of destinations served, as well as the frequency of those services. Two performance metrics are proposed in order to gauge the level to which this goal is achieved:

The first measure is the number of non-hub destinations served nonstop from any airport in the metropolitan area. This measure maps directly to the goal. Destinations which were served only on an occasional basis should not be considered and a lower bound of service at least once per week is imposed.

The second measure is the average daily frequency of service to the top domestic hubs (the definition of top domestic hubs is treated in section 4.3.2.4.2). This measure addresses the goal in two ways:

- It gives an indication of the level of frequency of service across a set of key routes
- It is a measure of the level of ease with which a large number of destinations can be reached through a single connection

These two measures reflect the two factors that impact total trip time, as discussed by (Belobaba et al. 2009, pp. 58-59). Total trip time involves both the time on board the aircraft as well as “schedule displacement,” with the latter being the amount of time that passes between a passenger’s desired departure time and the time when a flight is available. The number of destinations served nonstop will contribute toward minimizing the time on board the aircraft, and a high frequency of flights will minimize the schedule displacement.

4.3.2.2.2 Measuring Passenger Volume

The goal includes ensuring that the number of passengers using the airport is maximized. The goal can be measured in terms of passenger enplanements or deplanements, or the sum of the two. A distinction is made between connecting passengers and origin and destination (O&D) passengers. Whether or not it is desirable to maximize both passenger types is a multi-faceted question: For some stakeholders the passenger volumes should be maximized, irrespective of the passenger type.

For example, concessionaires and airport managers have an interest in ensuring that both passenger types are maximized since both types result in increased revenues. In contrast, for local communities, connecting passengers have both pros and cons: One drawback is that a large volume of connecting passengers means that congestion at the airport may go up due to increased traffic, causing inconvenience to local residents. In contrast, one benefit for local communities of connecting traffic is that the amount of air service at the airport can be maximized in the form of increased nonstop services and higher frequency; if passengers at the airport were purely O&D passengers, demand would be lower and as a result, less air service would be provided. A second benefit is that local businesses that operate at the airport and in surrounding areas will see their revenues go up as a result of connecting passengers, which provides increased employment and economic benefits to the regional community.

Based on this discussion, connecting passengers represent a desirable factor to a greater degree than they do a drawback and were included in the total passenger calculation.

4.3.2.2.3 Measuring Airport Capacity

Past studies of airport performance have measured airport capacity in terms of a number of different metrics of infrastructure size, as shown in Table 4.14.

Table 4.14 – Sample measures of airport capacity used in past benchmarks

Capacity Measure	Studies which use this metric
Number of runways	(Bazargan & Vasigh 2003) (Gillen & Lall 1997) (Sarkis & Talluri 2004) (Sarkis 2000) (Oum & Yu 2004)
Number of gates	(Bazargan & Vasigh 2003) (Gillen & Lall 1997) (Sarkis & Talluri 2004) (Sarkis 2000) (Oum & Yu 2004)
Terminal area	(Gillen & Lall 1997) (Pels et al. 2001) (Oum & Yu 2004) (Barros 2008)
Airport area	(Gillen & Lall 1997) (Pels et al. 2001) (Barros 2008)
Runway area/runway length	(Gillen & Lall 1997) (Pels et al. 2001) (Barros 2008) (Abbott & Wu 2002)
Number of aircraft remote and terminal parking positions	(Pels et al. 2001)

The primary factor which limits the amount of traffic that an airport can handle is the runway capacity. This, to a greater degree than any other factor, is the bottleneck in the airport system (Neufville & Odoni 2003, p. 367).

Airport runway capacity is not determined only by the number of runways but also by their geometric layout and by exogenous factors such as weather conditions (Neufville & Odoni 2003, p. 376). For example, two closely spaced runways (with centerlines less than 2,500 ft apart) will have a lower total capacity than two runways spaced further apart (Neufville & Odoni 2003, pp. 384-387), all other factors being the same. Similarly, a runway at an airport whose weather conditions permit visual meteorological conditions (VMC) operations more often will have a higher total capacity than one at an airport with better weather conditions (Neufville & Odoni 2003, p. 389).

Because of these differences in runway capacity, it is not sufficient to study airport capacity simply by counting the number of runways, as has been done in past studies. Instead, a measure of the actual capacity of the set of runways at an airport is necessary. An approach to determining actual airport capacity has been proposed (Kumar & Sherry 2009) in which airport Capacity Coverage Charts (CCCs) were used along with data on the costs of delay to determine average airport capacity. CCCs describe how much runway capacity was available and for how long (Neufville & Odoni 2003, p. 402). This average airport capacity is the measure used in this analysis.

4.3.2.3 Choice of Benchmark Model

The parameters for the model are the number of nonstop non-hub destinations served, the average daily frequency of service to the top domestic hubs, the total number of enplaned passengers, and the average airport capacity. This model can conceptually be expressed as the ratio (destinations served, frequency, passengers) : (capacity). The units of measure for these metrics are airports, flights, passengers, and aircraft, respectively. Combining these metrics into a comparative benchmark is a case where the analysis combines multiple parameters of different units, and where the production or utility function is unknown. As discussed in section 0, this makes DEA the appropriate benchmarking methodology.

The results of the application of the framework and heuristics to determine a model for this analysis are now presented.

- **Aggregation**: The heuristics specify that either an ε-maximin function or an additive function should be used. The additive function should be used only if a motivation exists for why the current proportional mix of inputs or outputs (depending on the orientation chosen) is irrelevant and can be changed. Otherwise, the ε-maximin function should be chosen. In this study, no evidence exists to suggest that the proportional mix of input or outputs can be changed between different airports. As a result, the ε-maximin function is chosen.

- **Weights**: The heuristics prescribe the use of specific weights unless any reasons are present for choosing range-adjusted weights. The specific weights allow each DMU to select its own optimal weights and in this study that is an appropriate selection to reflect the decisions of those involved in managing services at the airport.
- **Orientation**: The model orientation choice (input or output oriented) should be based on which among the model parameters are to a greater degree controllable by management. In this analysis, the output parameters are to a greater level possible to control or influence by the entities involved in managing the airport and its services. In contrast, the input in the form of runway capacity is largely a static value which is difficult to influence; once a runway has been constructed it is difficult to remove it (Martín & Román 2001, pp. 152-153), and conversely at some airports space constraints and community opposition limit the ability to add further runway capacity (Neufville & Odoni 2003, p. 168).

- **Returns to scale**: The framework specifies a choice between constant returns to scale (CRS) and variable returns to scale (VRS). The outputs in this model can both be assumed to reflect VRS: The number of new destinations which are feasible to serve decreases as the number of already served destinations increases, since only a finite number of metropolitan areas exist where the local market provides sufficient demand to warrant nonstop service.
- **FDH**: The Free Disposal Hull should be applied only if some reason exists why comparison only to observed combinations of inputs and outputs should be made, but no such reason exists in this analysis.
- **Integrality**: Integrality constraints should be applied in cases where input or outputs are indivisible into fractions and of low magnitude, and if large errors in the results would be introduced if these inputs or outputs were assumed to have decimal values. The parameter with integrality constraints and the lowest magnitude in this study is the number of non-hub destinations served nonstop, but with a median value of 88 for the years studied, this parameter's magnitude remains sufficiently high that no integrality constraints are necessary in the model. Although the runway capacity could also be subject to integer

constraints in its original form since the number of aircraft movements is an indivisible value, the adjustments necessary in this analysis to reduce the hourly capacity to only the portion used by domestic passenger traffic causes the capacity to take on non-integer values. As a result, no integer constraints can be placed on the runway capacity, causing some level of error in the results.

- **Timespan**: If any key technology changes have occurred during the timespan being studied that would impact the ability of DMUs to achieve strong performance, then a Malmquist index method should be used. If not, the performance for each year can simply be analyzed independently. In the present analysis, technology changes would involve the introduction of something which made it feasible for air carriers to serve more destinations than before, or something which allowed for increased frequency of service. From a technology point of view, this would involve the introduction of new aircraft types with highly different performance characteristics in terms of for instance fuel consumption, crew requirements, or number of seats. No new aircraft models for domestic use entered into service during the 2005-

2008 period from Boeing (The Boeing Company 2010), Airbus[7] (Airbus S.A.S. 2010), Bombardier (Bombardier 2010), or Embraer (Embraer 2010). As a result of no major changes occurring in this time period, no Malmquist index calculation is necessary.

- **Tie-breaking**: The heuristics prescribe that a tie-breaking function be used only if a reason exists why all airports must be fully ranked. No such reason is present and accordingly, no tie-breaking function is used.

Table 4.15 summarizes the modeling assumptions for this analysis.

Table 4.15 - DEA model parameter choices

Scalarizing function			Technology			Timespan	Tie-breaking
Aggre-gation	Weights	Orient-ation	Returns to scale	FDH	Integrality		
ε-maximin	Specific weights	Output oriented	VRS	No use of FDH	No integrality constraints	No use of Malmquist index; simply one analysis per year	None

[7] The Airbus A380 was in fact first delivered in 2007, but this aircraft is not used for US domestic service

These modeling assumptions are represented in the output-oriented BCC (R. D. Banker et al. 1984) algorithm with minimum weight constraints, which was used in this analysis. This model has the following dual problem formulation, as discussed further in section 2.2.2.6.2:

$$\max(\phi_a, \lambda) = \phi_a + \varepsilon\,(s^+ + s^-)$$

Subject to

$$\phi_a y_a - Y\lambda + s^+ = 0$$
$$X\lambda + s^- = x_a$$
$$e\lambda = 1$$
$$\lambda \geq 0, s^+ \geq 0, s^- \geq 0$$

The DEA scores were computed using the BCC implementation in Matlab, as discussed in section 3.5.1. For the implementation, the infinitesimal constant ε was set to 1.0 * E-6. A further discussion of the choice of this value is in section 4.3.3.6.1.

4.3.2.4 Data Collection and Pre-Processing

This section describes the means of obtaining and preparing the benchmark data for the analysis.

4.3.2.4.1 Data Sources

Two data sources were used for the analysis:

- **Data on destinations, frequencies, and passenger volumes**: This data was prepared using the T100 database which is compiled from data collected by Office of Airline Information (OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b). The T100 database is a complete census of flights by U.S. and foreign carriers and provides data on the number of operations and passengers carried between each city pair.
- **Data on airport capacity**: This data was derived from the analysis described in (Kumar & Sherry 2009). This analysis in turn was conducted using the Aviation System Performance Metrics (ASPM) database (Federal Aviation Administration 2010d) along with the T100 database described above and the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c).

4.3.2.4.2 Defining Hubs

The definition of all-points domestic hubs in the analysis was based on an initial analysis of the T100 database. The objective was to identify those airports that provide connections to the largest number of other airports. For the 2005-2008 time period, the analysis found the number of domestic airports served nonstop[8] presented in Table 4.16, and identified the average number of other OEP-35 airports served nonstop listed in Table 4.17.

[8] Only destinations that were served at least 52 times per year were considered, to ensure that at least weekly service existed.

Table 4.16 - Average number of domestic airports served nonstop at least 52 times annually (source: T100 database)

Airport	Average number of domestic airports served nonstop	Rank
ATL	171	1
ORD	141	2
DFW	138	3
MSP	137	4
DEN	134	5
DTW	128	6
IAH	121	7
LAS	119	8
CVG	119	9
CLT	102	10
SLC	101	11

Table 4.17 - Average number of OEP-35 airports served nonstop at least 52 times annually (source: T100 database)

Airport	Average number of OEP-35 airports served nonstop	Rank
ATL	34	1
DEN	34	1
DFW	34	1
MSP	34	1
CVG	33	5
DTW	33	5
IAH	33	5
LAS	33	5
LAX	33	5
ORD	33	5
PHX	33	5

The first four airports in Table 4.17 were connected to all other OEP-35 airports in each of the years from 2005 to 2008. In addition, these airports all rank among the top five airports in terms of the overall number of domestic destinations served, as shown in Table 4.16. The remaining top-five airport from Table 4.16 is ORD which, although it lacks service to one of the OEP-35 airports, ranks as the second most connected airport to other domestic airports. Based on this data, the list of hubs for this analysis is: ATL, ORD, DFW, MSP, and DEN. The impact of this definition is tested as part of the sensitivity analysis discussed in section 4.2.2.6.

4.3.2.4.3 Preparing Benchmark Data

Each of the data sources required some pre-processing for use in the benchmark analysis. This section describes that pre-processing.

The data on the number of non-hub destinations served nonstop was computed from data using these conditions and assumptions:

- Departures were considered from the metro area as a whole rather than from individual airports. For instance, if both EWR and LGA airports in the New York region had nonstop service to MSP, this would only be counted as one nonstop destination for the New York metropolitan area.
- At least 52 flights during the year were required in order for an O&D pair to be considered to have nonstop service.

The data on the daily frequency of service to hubs was prepared using these conditions and assumptions:

- As for the number of non-hub destinations served, departures were considered from the metro area as a whole rather than from individual airports. However, in the example with EWR and LGA above, if each airport had service four times daily, the New York region would be counted as having a frequency of eight.
- For those airports that were hubs, only service to the four other hubs could be counted while for non-hub airports, service to the five hubs was counted. To adjust for this, the hub airports' totals were increased by the average of their service to each of the other four hub airports; in practice this amounted to a multiplication of each hub airport's total by a factor of 1.25.

The data on the number of passengers carried was limited to only domestic passengers, in order to match the domestic-only scope for other performance parameters. The number of passengers could be counted as the number of enplaned passengers, the number of deplaned passengers, or as the sum of the two. The number of departing passengers is very close to the number of arriving passengers[9]; for convenience, the annual number of enplaned domestic passengers was chosen. In the database query, the measure was computed as the sum of all enplaned passengers for whom the origin was one of the OEP-35 airports.

Capacity data for airports was derived from (Kumar & Sherry 2009). This was the capacity for all operations but the scope of the study is only domestic passenger flights, the capacity value had to be reduced to only account for the portion of capacity used by domestic passenger flights. Any capacity used by international or cargo aviation, etc., had to be removed. The percentage of flights that were domestic passenger flights was computed by comparing the sum of all flights in the T100 segment database for all international and domestic flights with the sum of flights in the T100 domestic segment database that had a number of passengers greater than 0. The average percentage of flights that were domestic passenger flights across 2005-2008 is presented in Table 4.18.

---

[9] In a test of the number of domestic passengers for 2008 at the OEP-35 airports, the number of enplaned passengers was within 0.77% or less of the number of deplaned passengers.

The capacity used for domestic flights was computed by multiplying the original capacity by the portion of all flights that were domestic passenger flights.

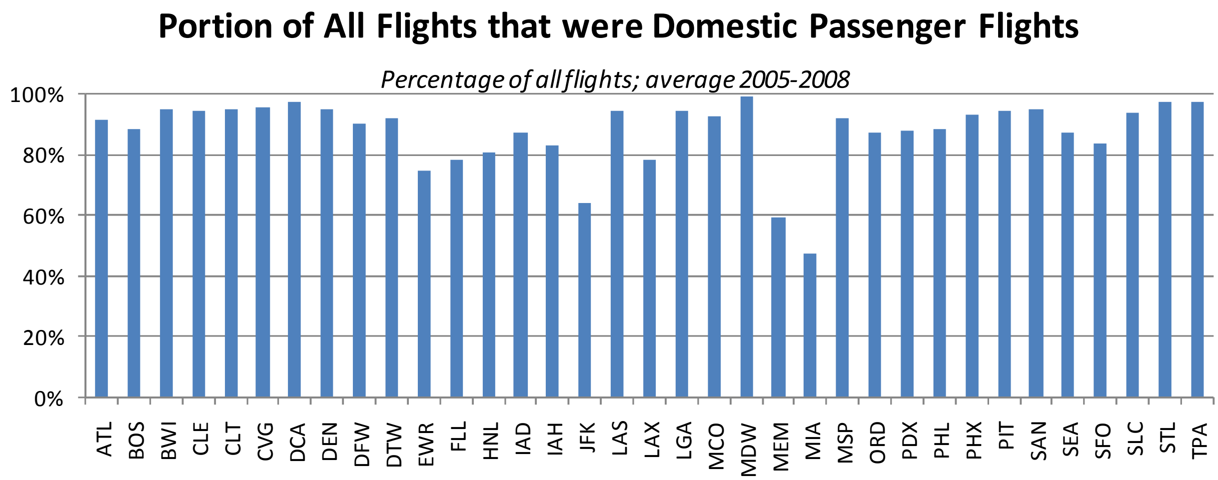

Table 4.18 - Portion of all flights that were domestic passenger flights

4.3.2.5 Summary of Input and Output Parameters

This section provides four-year average values for the input and three output parameters used in the DEA analysis. The full details of the input and output parameters are provided in Appendix E. Although the analysis was done separately for each of the four years, this overview provides averages for the whole period 2005-2008.

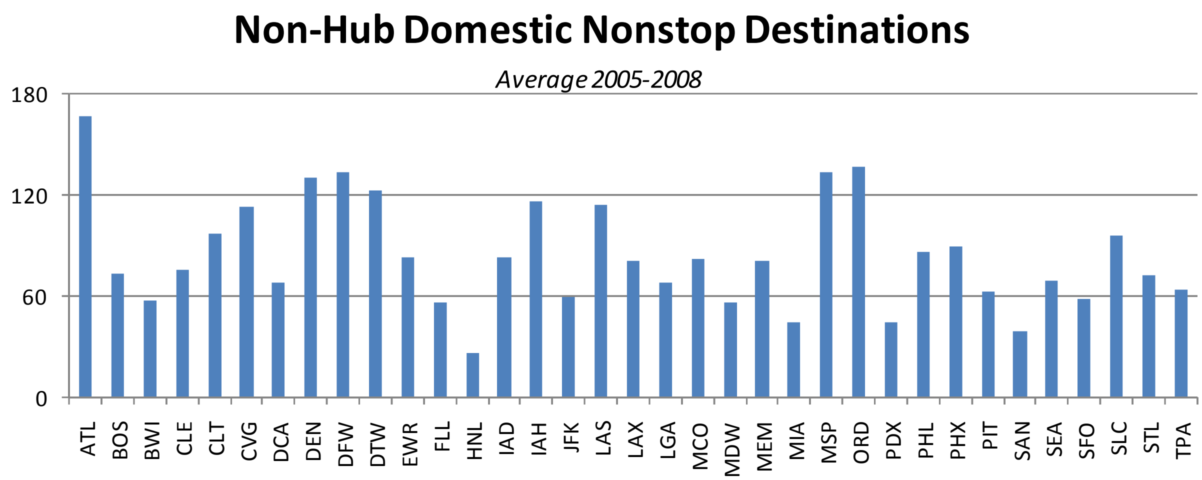

Figure 4.16 - Number of non-hub domestic destinations served nonstop, average 2005-2008

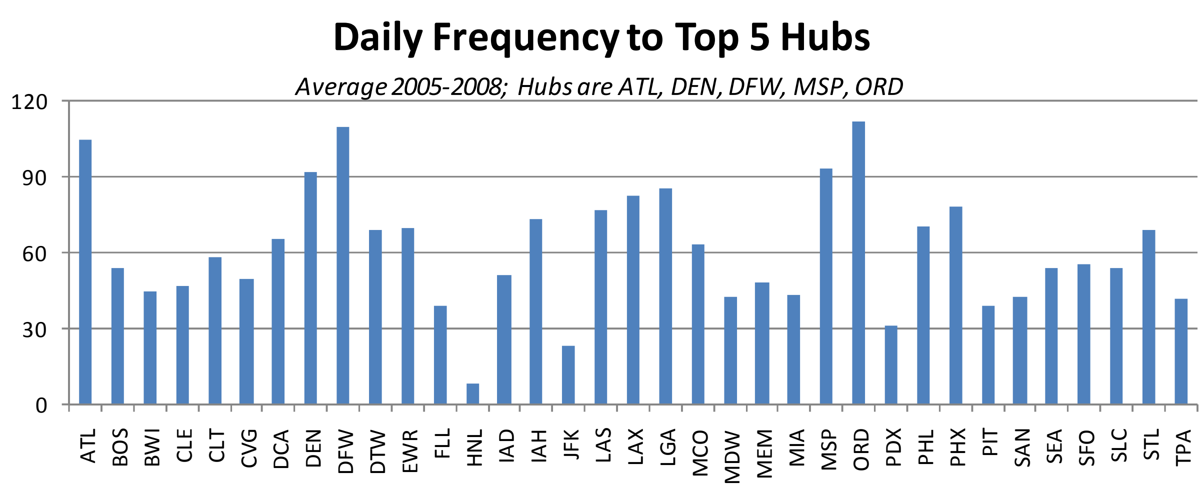

Figure 4.17 - Daily service frequency to top 5 hubs, average 2005-2008

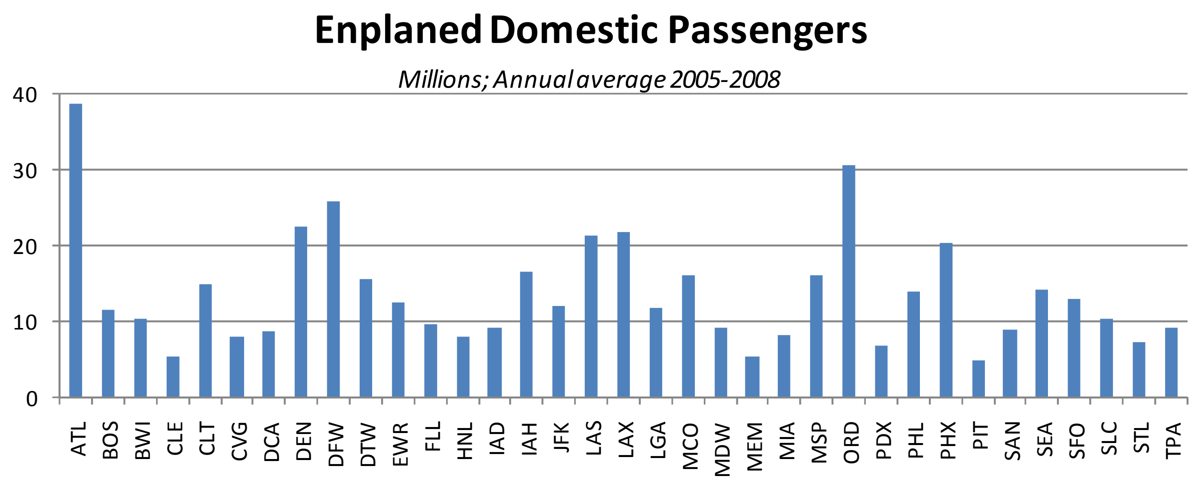

Figure 4.18 - Annual number of enplaned domestic passengers, average 2005-2008

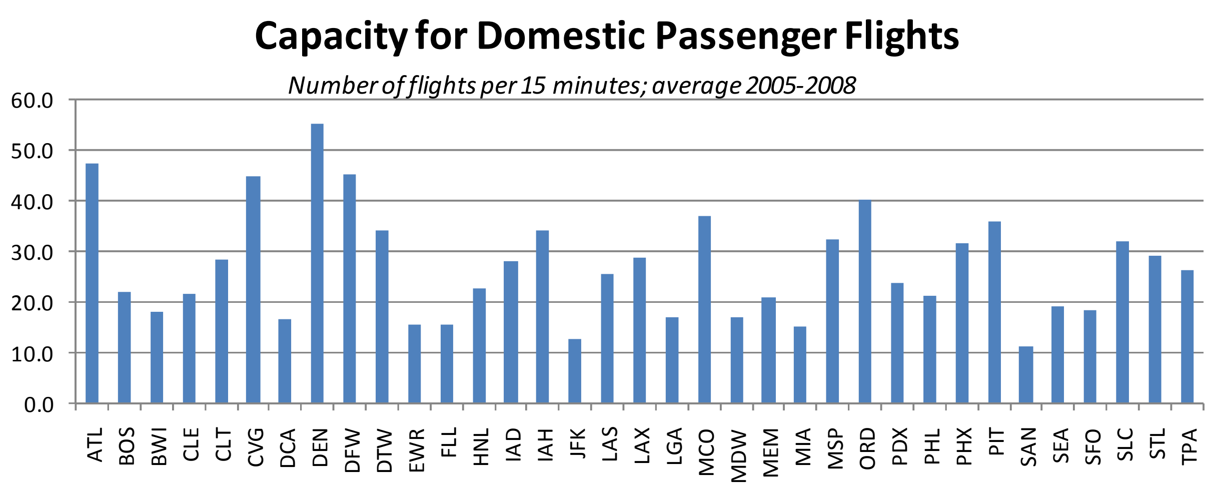

Figure 4.19 – Capacity for flights every 15 minutes, adjusted for domestic flights only, average 2005-2008

4.3.2.6 Sensitivity Analysis

The purpose of the sensitivity analysis is to understand the degree to which the findings stand up to any potential changes in the input and output data or the underlying model assumptions of the study.

The choice of DEA model has been shown to have a potentially radical impact on the results of airport performance studies (Schaar & Sherry 2008). Some studies have attempted to address that by using a variety of different models (Sarkis 2000), but this can lead to contradictory and inconclusive results. This paper instead used the framework and heuristics from section 0 to guide model selection. Any variations of the results based on using another DEA model would not be relevant since such a model would be selected without a rationale for its applicability. As a result, no sensitivity analysis using a different DEA model was conducted.

However, in the study of DEA models which use minimum weights, a large body of work exists (e.g. (Mehrabian et al. 2000) and (Allen et al. 1997)) but no conclusive determination of a standard approach to the choice of minimum weights exists. To address this lack of standardization, the sensitivity analysis in this study includes tests of varying these minimum weights.

The data on output parameters regarding the number of non-hub destinations served nonstop and the frequency of service to the top 5 hubs was based not on sampling data but rather on full census data. This means that no sensitivity analysis is necessary to test the impact of sampling errors. However, the data on both of these performance parameters is dependent on the definition of hubs. To test the robustness of the findings with respect to the definition of hubs, the sensitivity analysis included tests of using the top 3, 4, 6, and 7 hubs based on the total number of domestic destinations served nonstop (the list of these airports can be found in Table 4.16).

Regarding the total number of domestic passengers enplaned, no assumptions had to be made. Similarly, the data on the portion of airport capacity used for domestic traffic did not require any assumption other than that the portion of airport capacity used for domestic passenger traffic is proportional to the portion of airport traffic that is made up of domestic passenger flights. No sensitivity analysis of variations in enplanement data or airport capacity was conducted.

The results of the sensitivity analysis tests are presented in section 4.3.3.6.

4.3.3 Results

This section presents the overall results of the analysis and discusses several factors which impact the results. Figure 4.20 shows an average of the results of the benchmark of airport capacity utilization, with low values indicating strong utilization and high values indicating poor utilization. The full details of the results are provided in Appendix E.

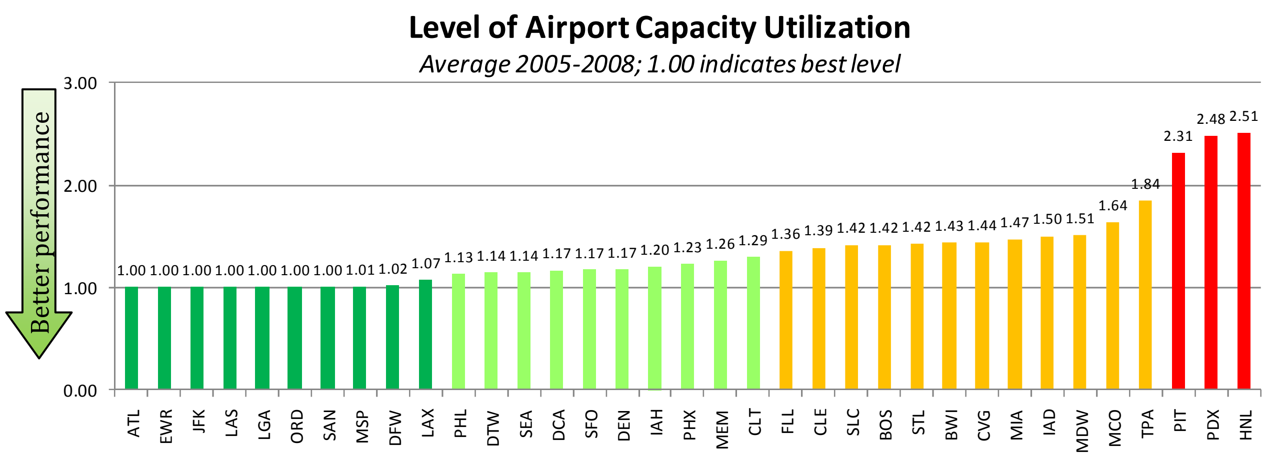

Figure 4.20 - Average level of airport capacity utilization performance. 1.00 indicates strong utilization and high values indicate poor utilization. Bar coloring is based on a k-means cluster analysis of benchmark results.

The results show a range of values that can be grouped into three categories:

1. Airports with very high utilization, shown as a score of at or near 1.00. These airports are a mix of the very large, such as ATL, ORD, and DFW; and the highly capacity constrained, such as SAN with only a single runway
2. Airports with mid-range levels of capacity utilization and scores ranging between 1.10 and 1.50.
3. Airports with poor utilization and scores above 1.50. Three airports in this category stand out with scores well above 2.00: PIT, PDX, and HNL. Such high values indicate that these airports' infrastructure capacities have potentially been over-built in relation to the actual demand for air services; in the case of PIT, the infrastructure capacity may have been designed with the intent of accommodating hub service, but with the loss of the US Airways hub operation at that airport, excess capacity exists.

The next several sections address factors which impact and are impacted by the level of capacity utilization.

4.3.3.1 Relationship between Delays and Capacity Utilization

Two different delay metrics relate to the level of airport capacity utilization. The first type of delay is the level of taxi-out tarmac delay. This is the time elapsed between the aircraft pushing back from the gate until wheels off, subtracting out the unimpeded taxi time (i.e. the amount of time the taxi would have taken, had there been no queue for reaching the runway). This delay relates in part to constraints on the runway capacity at the departure airport. The second type of delay is gate arrival delay. This is the difference between the scheduled arrival time and the actual arrival time. This measure is indicative of the constraints on runway capacity at the arrival airport.

Other types of delay also exist in the form of taxi-in delay and gate departure delay. Taxi-in delay is less frequently occurring[10] than taxi-out delay and does not relate to runway capacity but rather to tarmac configuration and gate availability. Gate departure delay similarly has less of a relationship to actual runway capacity

[10] Across all of the OEP-35 airports for the period 2005-2008, the average taxi-in delay was 2.34 minutes, while the taxi-out delay was 6.19 minutes (Federal Aviation Administration 2010c)

and is more driven by other factors causing airline delay. These types of delays are not reviewed in this analysis.

4.3.3.1.1 Taxi-Out Delay

The taxi-out delay is computed by measuring the difference between the actual gate out time and the actual wheels off time, and subtracting the unimpeded taxi time. The unimpeded taxi time is measured for each carrier in optimal conditions with no congestion, weather, or other delay factors present (Federal Aviation Administration 2010b).

A comparison of the level of capacity utilization and the average taxi-out delay is shown in Figure 4.21. The data reveals that although some airports achieve very high capacity utilization, it can come at a high taxi-out delay costs. The three New York area airports exhibit this combination. In contrast, airports such as MSP, LAS, DFW, and SAN exhibit high degrees of capacity utilization but do not show the same negative effects in the form of high taxi-out delay costs. Lastly, the data shows that the airports with the lowest degrees of capacity utilization also show some of the lowest levels of taxi-out delays, further suggesting that these airports have excess airport infrastructure capacity.

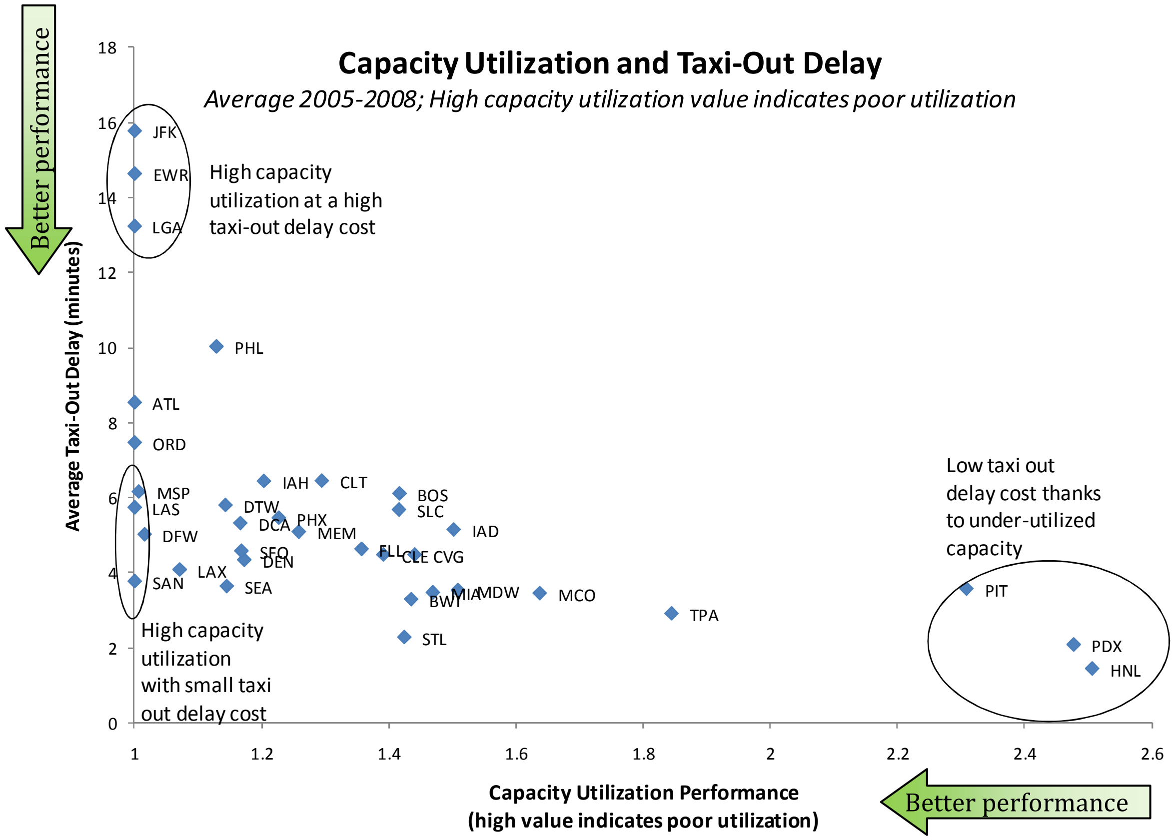

Figure 4.21 - Capacity utilization performance in relation to average taxi-out delays. A high capacity utilization value indicates poor performance

4.3.3.1.2 Gate Arrival Delay

The gate arrival delay is computed by calculating the difference between the scheduled gate in time and the actual gate in time. Unlike the taxi-out delay measure, the gate arrival delay can be compensated for by air carriers through schedule padding, a practice in which airlines add time to a flight's schedule in anticipation of delays (Long et al. 1999, pp. 2-8). The data on gate arrival delay in relation to the capacity utilization performance is presented in Figure 4.22 and shows a pattern similar to that of taxi-out delays. The New York area airports achieve high degrees of capacity utilization but operators in that market pay a price of high gate arrival delays. In contrast, several airports achieve high capacity utilization while maintaining lower levels of gate arrival delays, and the same three airports as in the previous section exhibit low gate arrival delays in part thanks to the availability of excess capacity.

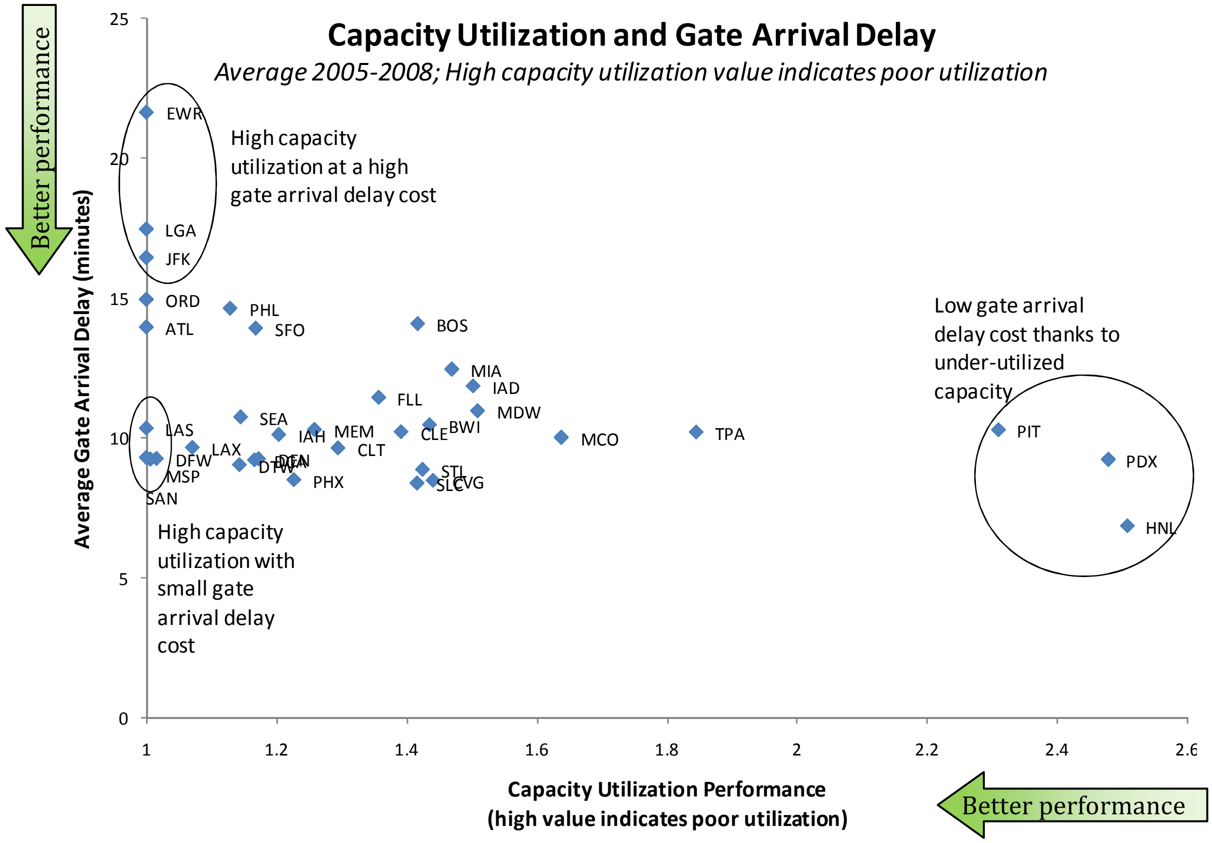

Figure 4.22 - Capacity utilization in relation to gate arrival delay. A high capacity utilization value indicates poor utilization.

4.3.3.2 Drivers of Passenger Volume

The volume of enplaned passengers is one of the factors that contribute toward the degree of capacity utilization at an airport. Assuming the presence of sufficient demand, the number of passengers carried is driven by the number of flights at an airport in relationship to the available capacity as well as by the number of seats on those aircraft. The degree to which that available capacity is actually used is expressed in the form of load factors, which is the portion of available seats that is occupied by passengers. Each of these factors is presented in this section.

The average of the annual number of domestic passenger flights per unit of airport capacity is presented in Figure 4.23. This data shows that three of the four proportionately busiest airports are located in the New York area; these airports also represent the highest levels of delay, as shown in the previous section. In contrast, the three airports with the lowest levels of capacity utilization rank among the five bottom airports on this measure.

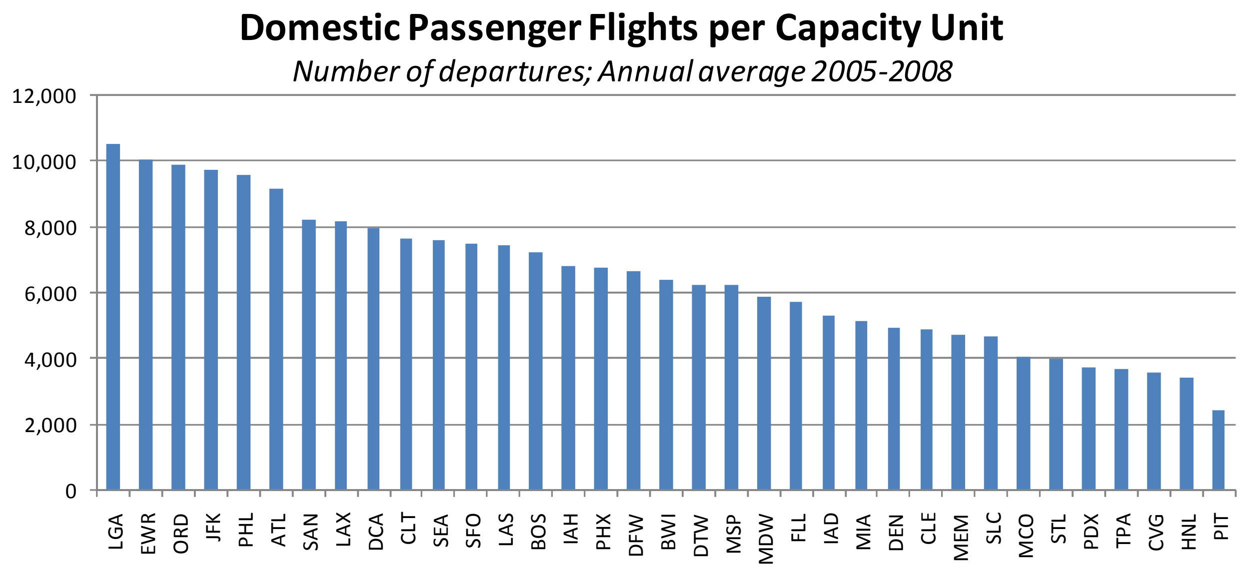

Figure 4.23 - Number of annual domestic flights per airport capacity unit; average 2005-2008

The average number of seats per flight is shown in Figure 4.24. This measure is also referred to as aircraft gauge. The data indicates that the highest-gauge markets are those that may be considered leisure destinations to a greater degree than others: LAS, FLL, MCO, MIA, and HNL. Among the five airports that exhibit the largest number of flights per available capacity unit in Figure 4.23, four fall in the lower half in terms of the number of seats per flight.

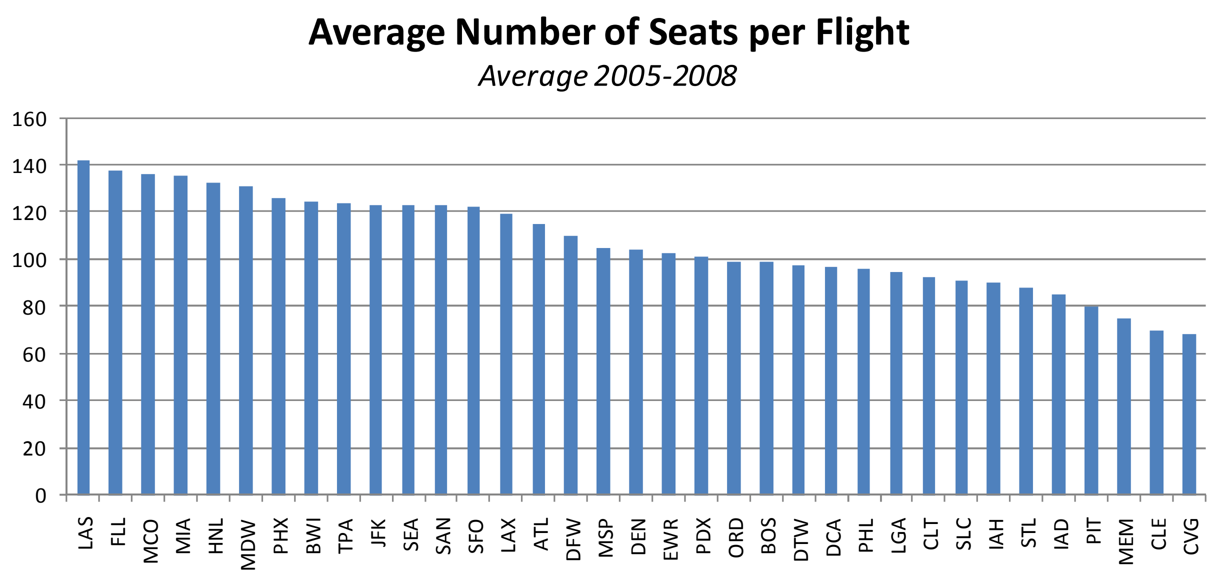

Figure 4.24 - Average number of seats per flight (also known as gauge), 2005-2008

Figure 4.25 shows the average load factor for each airport. This load factor is computed by summing the number of enplaned domestic passengers and dividing by the total number of seats on flights departing from the airport. Although other calculations are also used for the load factor which incorporate factors such as the distance traveled (American Airlines 2010), this analysis was based on the simpler measure since the primary measure of interest is the number of seats out of the airport that were occupied. Airports that rank at the bottom for this measure include LGA and DCA which are important shuttle markets, in which the airlines schedule departures with very regular intervals (often hourly) with the objective of capturing large portions of the business travel market (Plunkett 2007).

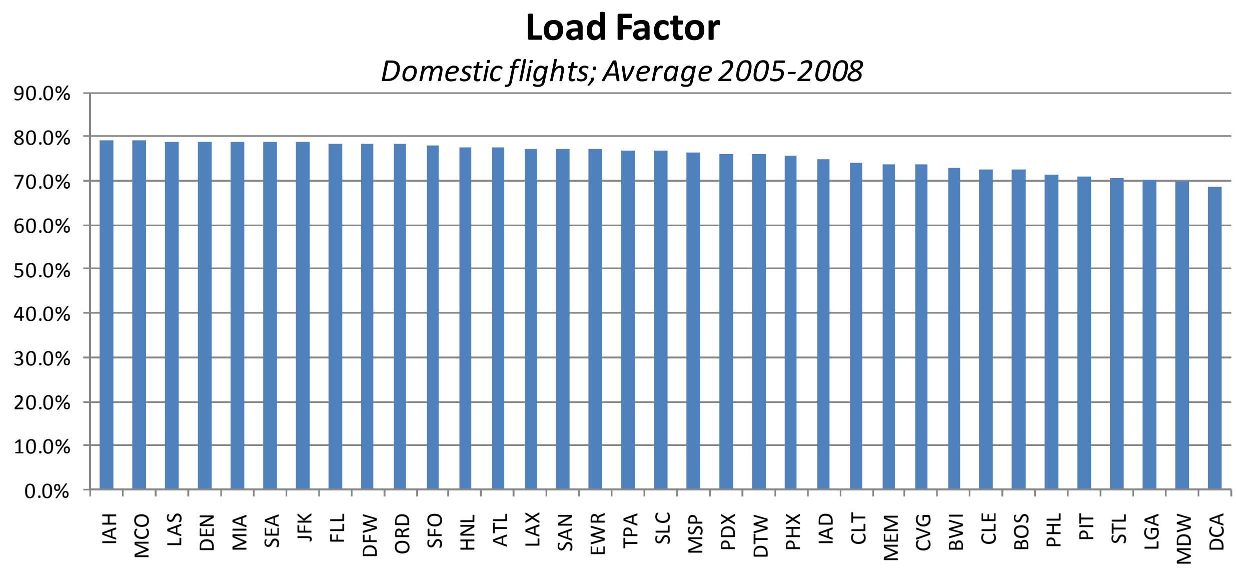

Figure 4.25 - Average load factor for departing domestic flights, 2005-2008. Although load factors can be computed in different ways, this simplified measure was computed by summing the number of enplaned passengers and dividing by the total number of seats on departing flights

Taking these factors together shows that the main contributing factor to ensuring high capacity utilization is a high number of domestic flights by airport capacity unit. However, for some airports that are already operating at a very high number of flights per capacity unit, delay data suggests that increasing the number of flights further would come at a cost in the form of increased delays. For those airports, increased passenger volumes could be achieved by increasing the average number of seats per aircraft, so-called "upgauging". However, upgauging requires that demand exists that will occupy these seats. For an airport like LGA, this may not be possible if current flight schedules remain un-altered, as indicated by that airport's comparatively low load factor.

4.3.3.3 Drivers of Level of Air Service

Just as the number of passengers, the number of nonstop destinations served and the frequency of service to hubs can be maximized when the number of flights in relation to the airport capacity is maximized. This measure is displayed in Figure 4.23.

To maximize the number of flights there must be an underlying demand for air travel in the region where the airport is located. This was addressed in case study 1 in section 4.2. The study used the same definition of air service as this paper and studied the same time period. Figure 4.26 shows the relationship between the level of regional air service and the level of capacity utilization.

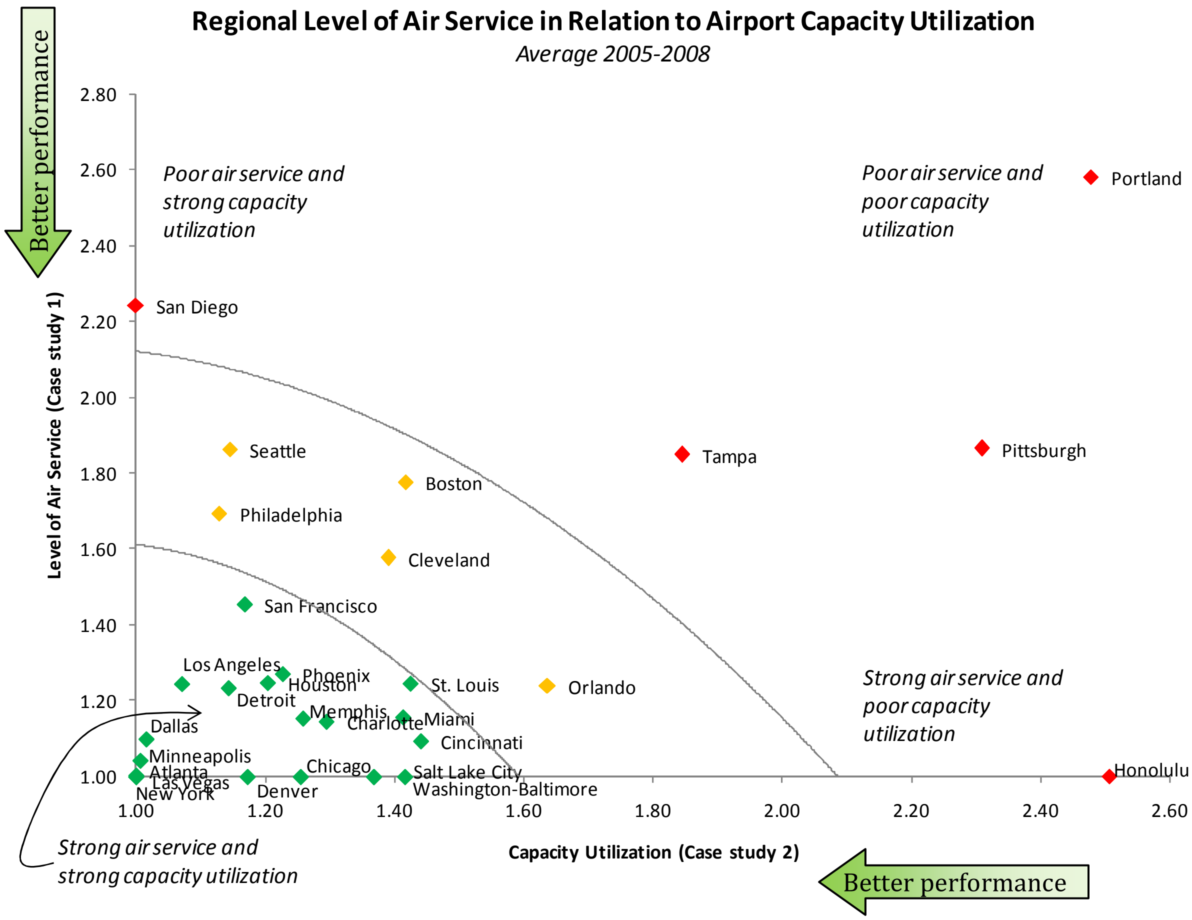

Figure 4.26 - Regional level of air service in relation to airport capacity utilization. For both values, 1.00 indicates best performance and the highest values represent the worst performance.

In Figure 4.26 the average level of capacity utilization was used for those areas that encompass more than one OEP-35 airport.

This data shows that some metropolitan areas such as New York report high levels of air service along with a high level of capacity utilization. This may be considered a cautionary data point for those regions because it indicates that if further population and economic growth occurs, an already high capacity utilization level may impose limits on the ability for the region's airports to accommodate further growth.

In contrast, those airports that have low levels of capacity utilization and exist in areas where the relative level of air service is already quite high can expect to continue experience low utilization levels since there do not appear to be the necessary population economic conditions for increased levels of air service. Honolulu is a particularly strong example in this category.

Finally, in areas with poor levels of air service but high airport capacity utilization, it appears that the reason for that poor level of air service is the lack of airport capacity. For instance, SAN is the busiest single-runway airport in the United States (San Diego International Airport 2010), and this data suggests that SAN could expect to see improved air service if it were possible to expand the airport's runway capacity.

4.3.3.4 Impact of Being a Hub

Hub- and-spoke carriers designate hubs so that connecting services can be provided to as many points as possible through connections at one or more hubs (Belobaba et al. 2009, p. 163). Airports that serve as hubs and carry high volumes of connecting passengers are able to achieve higher levels of air service than they would if they recorded primarily O&D traffic.

Data on the level of domestic O&D service at each of the airports was computed from the DB1B database (Bureau of Transportation Statistics 2010c) by summing all passenger itineraries that started or ended at an airport and dividing it by the sum of all enplanements and deplanements at that airport (including both connecting and O&D passengers).

Figure 4.27 shows that airports with high degrees of capacity utilization represent a mix of primarily O&D service airports (such as SAN and LGA) and airports with primarily connecting traffic (such as ATL and DFW). The data also shows that the airports with the lowest levels of capacity utilization – HNL, PDX, PIT, TPA, and MCO – are all primarily O&D service airports.

Increasing the level of connecting traffic by attracting a hub carrier would improve the level of capacity utilization at the under-utilized airports. However, the cause of the level of over-capacity at PIT can be traced in part to its loss of hub status in the US Airways network (Grossman 2007), suggesting that challenges exist in attracting new hub service.

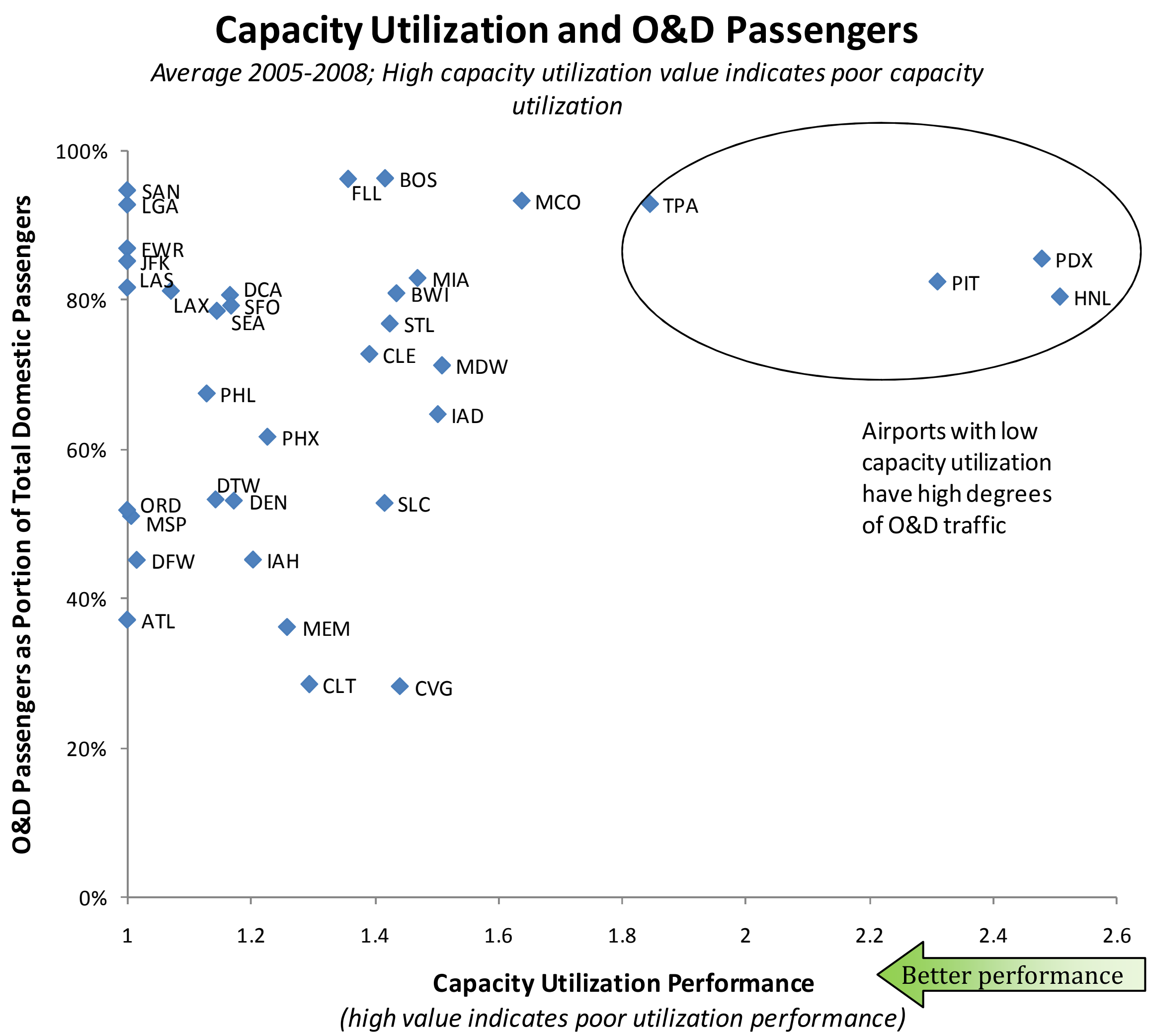

Figure 4.27 - Level of capacity utilization in relation to the portion of domestic passengers which are O&D passengers. Average 2005-2008. A high capacity utilization value indicates poor capacity utilization.

4.3.3.5 Gaps to Close

The underutilized airports are defined as those with scores greater than 1.00, and are considered inefficient in the DEA analysis. The DEA algorithm provides targets which DMUs should hit in order to move from inefficiency to efficiency. The targets are computed by multiplying each output by the DMU's efficiency score from the DEA analysis. These points are the projections for the DMUs on the convex hull represented by the efficient frontier.

These projections can provide improvement goals for managers at inefficient airports. When the original parameter values are subtracted from these targets, the gap that must be closed is obtained. Those gaps are presented in Table 4.19. The airports in Table 4.19 that have blank values for the gaps for all inputs are fully efficient in that year.

Table 4.19 - Distance to the capacity utilization frontier. These are gaps in the outputs to be closed for achieving a score of 1.00. The gaps are the distance to the frontier.

	Distance to Frontier											
	2005			2006			2007			2008		
	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax
ATL												
BOS	22.4	28.9	4.6	23.8	30.9	5.1	22.4	31.5	4.9	21.6	30.7	4.5
BWI	19.5	24.5	4.1	19.3	23.2	4.4	19.2	26.5	4.8	19.0	26.3	4.7
CLE	21.2	30.7	2.3	18.9	28.3	2.1	22.3	35.7	2.7	11.3	22.4	1.4
CLT	18.0	28.7	4.1	17.5	28.5	4.2	21.1	35.2	5.7	12.0	21.1	3.3
CVG	12.4	27.8	2.4	21.1	51.1	3.4	27.5	62.7	4.3	23.9	54.0	3.2
DCA	8.8	9.0	1.1	10.6	10.6	1.4	11.9	12.9	1.7	12.1	12.3	1.6
DEN	19.0	24.5	4.1	17.3	23.2	4.1	14.5	21.3	3.7	12.4	20.0	3.4
DFW	0.7	0.9	0.2	1.2	1.4	0.3	2.4	2.9	0.6	2.4	3.1	0.6
DTW	9.7	18.1	2.3	10.0	16.7	2.2	12.5	22.0	2.8	7.3	13.2	1.6
EWR												
FLL	11.3	16.9	2.9	15.6	20.6	3.8	16.1	21.3	3.9	12.4	20.5	2.9
HNL	12.4	37.8	10.9	11.4	40.9	11.5	10.6	38.9	11.8	13.2	42.7	12.9
IAD	30.6	39.6	5.1	26.1	40.9	4.4	27.2	47.8	5.3	19.0	36.9	3.7
IAH	15.6	23.9	3.2	13.8	21.5	3.1	16.6	26.2	3.9	13.6	22.5	3.2
JFK												

	Distance to Frontier											
	2005			2006			2007			2008		
	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax	Depts. to top 5 hubs	Non-hub non-stops	Num. pax
LAS												
LAX	5.2	4.8	1.4	7.3	7.2	2.0	4.6	4.6	1.2	6.1	6.2	1.6
LGA												
MCO	41.4	54.0	10.3	40.7	53.6	10.4	39.7	50.2	10.1	40.0	51.8	10.2
MDW	19.4	26.1	3.9	23.1	25.2	4.4	19.4	28.5	4.5	23.3	33.1	5.6
MEM	10.4	17.5	1.2	9.9	16.2	1.1	16.1	26.1	1.8	13.4	23.2	1.5
MIA	20.2	19.8	3.6	20.2	20.9	4.0	21.0	21.7	4.2	19.1	20.6	3.8
MSP							2.3	3.3	0.4			
ORD												
PDX	47.6	63.2	10.0	48.7	63.8	10.1	46.5	70.2	10.8	41.5	65.5	9.5
PHL	6.2	7.2	1.1	8.6	10.7	1.7	12.4	15.6	2.6	8.5	10.9	1.7
PHX	18.0	21.5	4.9	17.5	19.1	4.4	16.0	18.3	4.2	18.7	21.9	4.9
PIT	38.7	67.4	4.5	41.0	73.5	5.3	50.3	84.2	6.4	68.2	83.1	8.5
SAN												
SEA	7.4	9.6	2.0	9.2	11.4	2.3	8.6	11.0	2.3	6.1	7.8	1.6
SFO	9.1	9.3	2.0	12.2	12.4	2.7	10.0	10.4	2.3	6.3	7.2	1.6
SLC	28.2	50.1	6.0	17.9	32.2	3.3	20.1	37.6	3.9	22.6	38.2	4.0
STL	26.6	29.8	2.7	31.6	35.1	3.4	29.9	30.0	3.2	28.3	27.8	2.9
TPA	30.3	48.9	6.7	34.3	53.6	7.8	40.4	58.1	8.9	35.0	54.7	7.6

4.3.3.6 Results of Sensitivity Analysis

This section presents the findings from the sensitivity analysis that was described in section 4.2.2.6.

4.3.3.6.1 Sensitivity to Weight Boundaries

The sensitivity analysis applied different lower boundaries to the output weights to determine their impact on the capacity utilization scores and the rankings among airports.

In the original analysis, the standard weight boundaries ε from the BCC model were used. These are the boundaries on the minimum values on the weights applied to each output in the DEA calculation. In the BCC model these are simply specified as infinitesimal and in the model implementation, they were set at 1.0 * E-6.

In the sensitivity analysis, the weights were varied between the minimum value of 1.0 * E-6 up to the maximum feasible output weight values. The maximum feasible values are the maximum observed values multiplied by 1/3 (as a result of there being three output parameters). The maximum feasible values are those which result in all the constraints being binding for one or more DMUs.

The input parameter weights are not varied since any minimum values unfairly penalize the performance of the larger airports due to the differences in magnitude of the airports' values.

In the case where the analysis uses the maximum feasible weights, the DMU(s) with the highest magnitude of outputs are forced to apply exactly those weights, effectively removing the DMU's ability to select its own optimal weights. The higher the boundary on weights, the lower the flexibility for DMU's to determine their own optimal weights.

For the output weights, seven variations on the weight boundaries were tested for each year; the first test i=1 used the standard 1.0 E-6 weights, and in each subsequent test i=2..7 the boundary was proportionally increased such that the test i=7 had the maximum feasible boundaries (for tests i=2..7 the weight boundaries were determined as $boundary_i$ = max(weight) / 2 * (i – 1) / 6).

The average scores computed in the sensitivity analysis are presented in Figure 4.28. A comparison of the rankings of each metro area's scores between Test 1, Test 2, and Test 7 is presented in Table 4.20.

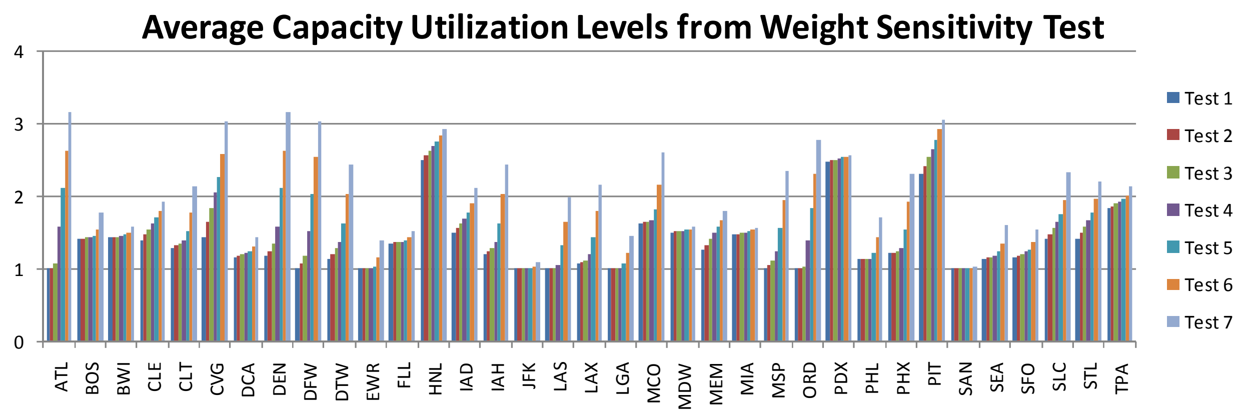

Figure 4.28 - Results from weight boundary sensitivity tests. Test 1 has the least restrictive weight boundaries, and Test 7 has the most restrictive boundaries

Table 4.20 - Rankings from selected sensitivity tests. Test 1 has the least restrictive weight boundaries and Test 7 has the most restrictive boundaries

	Ranking in Sensitivity Test			
	Test 1	**Test 2**	**...**	**Test 7**
ATL	1	1		34
BOS	24	22		13
BWI	26	23		10
CLE	22	24		15
CLT	20	19		19
CVG	27	31		31
DCA	14	13		4

	Ranking in Sensitivity Test			
	Test 1	Test 2	...	Test 7
DEN	16	18		34
DFW	9	9		32
DTW	12	15		26
EWR	1	1		3
FLL	21	21		6
HNL	35	35		30
IAD	29	29		17
IAH	17	17		25
JFK	1	1		2
LAS	1	1		16
LAX	10	10		20
LGA	1	1		5
MCO	31	30		28
MDW	30	28		9
MEM	19	20		14
MIA	28	25		8
MSP	8	8		24
ORD	1	1		29

	Ranking in Sensitivity Test			
	Test 1	Test 2	...	Test 7
PDX	34	34		27
PHL	11	11		12
PHX	18	16		22
PIT	33	33		33
SAN	1	1		1
SEA	13	12		11
SFO	15	14		7
SLC	23	26		23
STL	25	27		21
TPA	32	32		18

The sensitivity analysis shows that the rankings between Tests 1 (which is the case used in the full study) and Test 2 remain relatively unchanged. These are the tests which use the lowest boundaries. Among the 35 airports, 18 do not change their rank between the two tests. 12 airports move one or two rank levels, and 5 airports move three ranks or more. As the tests continue and the weights become more restrictive, the rankings change more strongly, and by Test 7, the results are radically different from the original results.

This shows that the selection of weight boundaries do matter to the results if they go well above the infinitesimal. However, the BCC model specifies that infinitesimal weight boundaries be used, and the similarity between the results of Test 1 and Test 2 shows that the exact choice of infinitesimal weight boundaries in the model implementation has little impact; the boundaries in Test 2 already far exceed what could be considered reasonable infinitesimal weight boundaries in the model. This indicates that the boundaries of 1.0 * E-6 used in the analysis are acceptable.

4.3.3.6.2 Sensitivity to Hub Definition

In the sensitivity test where the definition of hubs was changed as described in section 4.2.2.6, tests were run for 3, 4, 5, 6, and 7 hubs. The results were then averaged across all cases, and standard deviations for the level of air service were computed. The results of this analysis are shown in Figure 4.29.

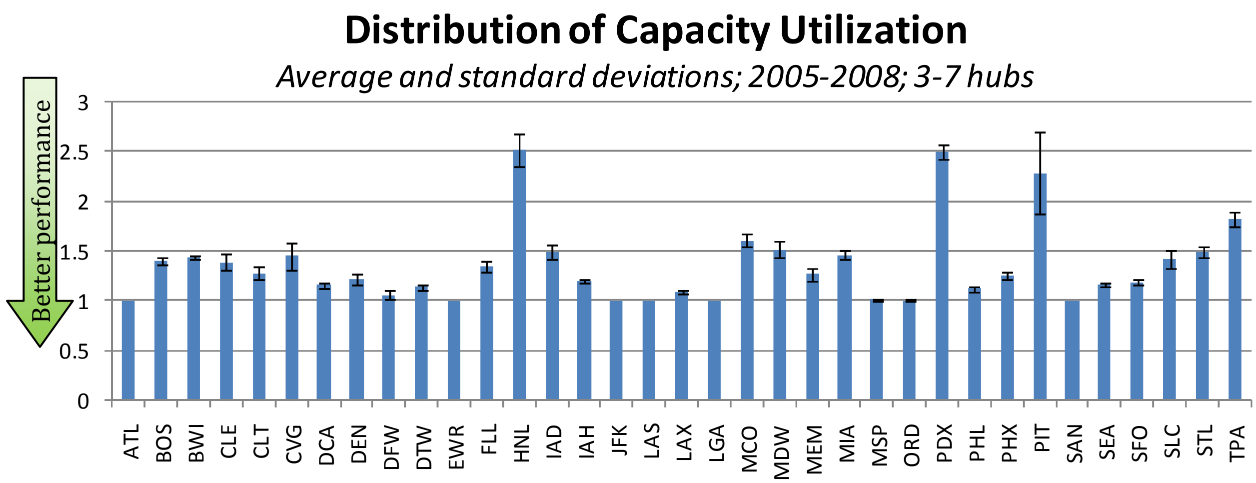

Figure 4.29 - Results from hub definition sensitivity test. The bars show the average score, and the error bars show +- the standard deviation

The results show a very small standard deviation for the fully efficient airports and limited standard deviations elsewhere, indicating a limited impact on the results from changes in the definition of how hubs are determined. The main exception to this is PIT, which does exhibit a higher standard deviation. However, the reason for PIT's high standard deviation stems not so much from changes in the hub definitions but rather from the fact that PIT's score worsened for each year of the analysis; for the period 2005-2008 the annual averages across the different hub definitions were 1.87, 2.06, 2.27, and 2.90.

4.3.3.7 Study Limitations

Although the extent of their impact on the study's results is unknown, some limitations to the results exist:

- The calculation of the level of air service does not factor in the geographic location of the metropolitan area. It is possible that areas located near the center of the continental United States have an inherently greater possibility of achieving high levels of air service.
- The calculation does not consider the relatively close proximity of some metropolitan areas to other areas. It is possible that the proximity to another area impacts a region's level of air service.
- International traffic was excluded from the study since 14 airports among the OEP-35 airports represented 70% of all international passenger enplanements in 2006 (FAA 2008, pp. 23-24). A study that included international traffic would show different results.
- Although the runway capacity is a value which would originally take on only integer values, the preprocessing necessary to incorporate only the portion of capacity used by domestic passenger flights causes this value to take on non-integer values. This introduces an error in the results, the impact of is not quantified.

4.3.3.8 Relationship with Capacity Utilization Benchmark Results and the Traditional Capacity Utilization Measure

The new measure of capacity utilization presented in this case study, reflecting the interests of several stakeholders, is presented as a contrast to the traditional measure of capacity utilization, as defined by the number of aircraft movements per unit of runway capacity. A high degree of correlation between the new metric and the traditional metric of capacity utilization would suggest that the traditional metric is a strong proxy of the degree to which stakeholders' interests are met; a lack of correlation would suggest that the traditional metric is a poor performance indicator.

To test this relationship, the relationship between the new metric and traditional metric of capacity utilization was visualized in a scatter plot, as shown in Figure 4.30, and a correlation test was conducted.

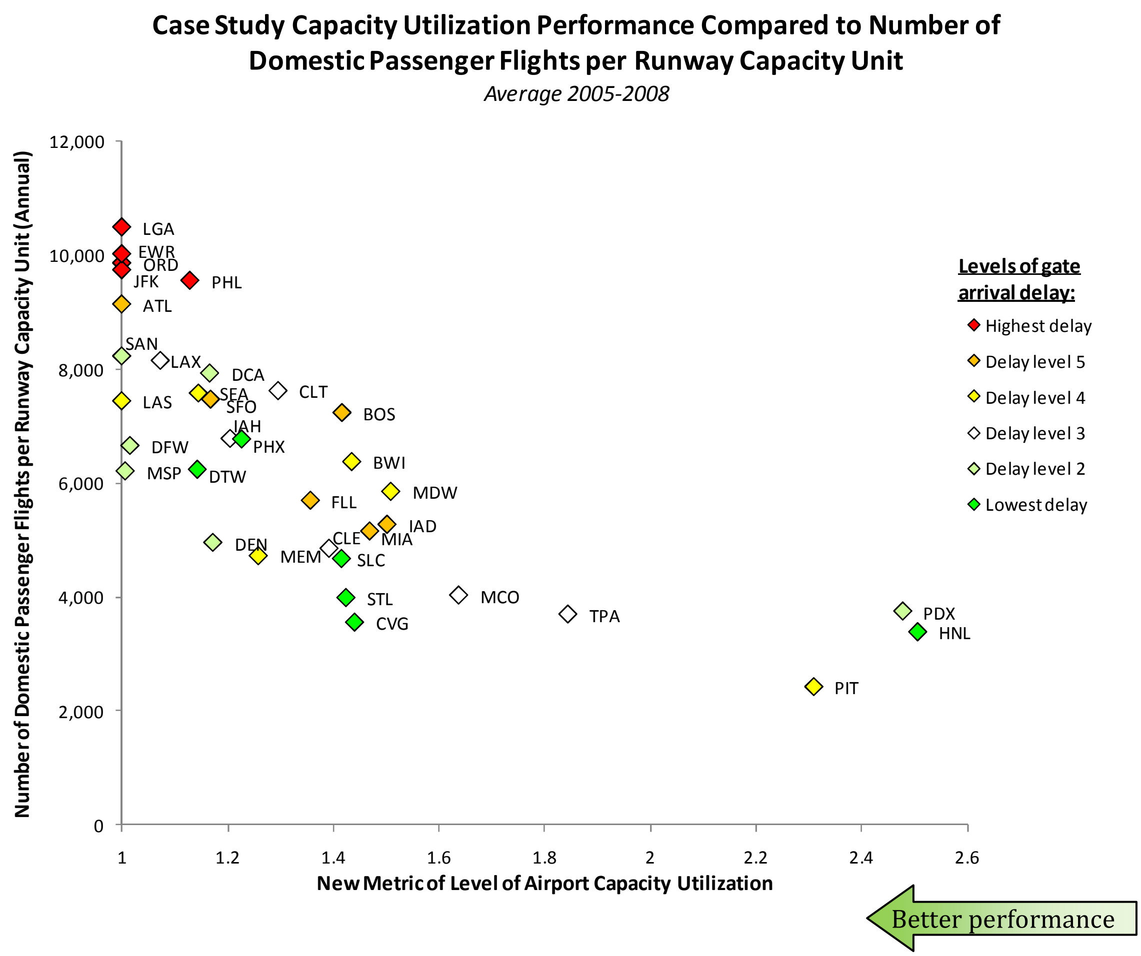

Figure 4.30 - Relationship between new metric (x-axis) and traditional metric (y-axis) of capacity utilization. Colors of diamonds indicate the levels of gate arrival delay.

The scatter plot suggests that a relationship exists, and the correlation coefficient between the two metrics is -0.75 ($p=1.86 * 10^{-7}$). The negative sign of the correlation coefficient reflects the fact that the benchmark was conducted in output-oriented mode, resulting in high capacity utilization being described with a low value, and vice versa. In contrast, for the traditional measure, high utilization is indicated by a high value. The high magnitude of the correlation coefficient indicates that a strong relationship exists, which indicates that the traditional measure of capacity utilization is a useful indicator and a strong proxy for the degree to which stakeholders' interests are being met.

In the chart is also displayed the level of gate arrival delay at each airport, with six equally sized groups of airports created based on the level of gate arrival delay. The chart indicates that all airports with the highest level of traditional capacity utilization belong to the group of airports with the highest levels of gate arrival delay. In contrast, airports such as SAN, LAS, and MSP are deemed fully efficient in the new capacity utilization benchmark and achieve that rating without incurring the highest levels of delay.

4.3.3.9 Summary of Capacity Utilization Performance

The analysis identified several categories of airports in terms of capacity utilization performance. The first group is that of the fully efficient airports, which in turn can be sub-grouped into two different categories:

The first sub-category is fully efficient airports at which operators pay a high price in the form of both taxi-out and gate arrival delay. These airports include the three New York City area airports, EWR, LGA, and JFK. At these airports, adding more flights to continue improving the level of capacity utilization as evidenced by the fact that they represent three of the four airports with the highest number of flights per capacity unit; adding more flights is likely to result in further exacerbating the delay problem.

Instead, growth in capacity utilization at these airports must come from increasing the number of passengers carried, which would be possible if air carriers were to increase the size of the aircraft used; for instance LGA currently ranks as number 25 out of 35 for the average number of seats per aircraft. However, the solution is not simply to begin flying larger aircraft; there must also be demand to fill the available seats, and LGA currently ranks third from the end on that measure. This suggests that air carriers would also have to begin flying different routes where demand is higher. However, this may not be attractive to air carriers since it is likely that the routes that are currently being served are high-yield, if low-demand, routes. Routes with higher demand may not represent the same level of yield.

The second sub-category of fully efficient airports is those which have low levels of delay. They include DFW, MSP, and LAS. At these airports, continued growth in capacity utilization appears feasible without corresponding increases in delay. SAN is also a member of this group of airports with low delay levels but high capacity utilization. However, considering the high level of aircraft per unit of runway capacity, SAN appears to be at risk of experiencing increased delay levels if its level of air service expands.

At the opposite end of the spectrum are those airports with poor levels of capacity utilization. These are airports that have more capacity than is necessary for delivering the current levels of air service and passenger volumes. Within this group there are also two sub-categories.

The first sub-category is those airports that have poor levels of capacity utilization but have the potential for improvements in utilization. This group includes airports such as PDX, PIT, and TPA, where the current level of air service is poor. These airports exist in relatively under-served markets where conditions may permit increased levels of air service, which would in turn increase the level of capacity utilization.

The second sub-category is those airports that have poor levels of capacity utilization but high levels of air service. HNL is the strongest example in this category, and can also be categorized as a member. For these airports, the conditions necessary for improved levels of air service do not appear to exist, suggesting that the current poor levels of capacity utilization will persist.

Finally, for many of the airports, regardless of capacity utilization, any changes in hub service will have an impact on the level of capacity utilization. For the under-utilized airports such at PIT, PDX, and MCO, addition of hub service is likely to improve the level of capacity utilization. Conversely, for high-utilization airports with high levels of connecting passengers such as DFW, MSP, and ATL, any reduction or loss of hub service represents the risk of considerable worsening of the level of air service.

4.3.4 Conclusions

The analysis ranked the level of capacity utilization at the OEP-35 airports based on stakeholder goals and found a large number operating at strong levels of capacity utilization. However, impediments to continued growth in capacity utilization were identified for some airports in the form of delay costs. While increasing the size of aircraft used would appear to provide the opportunity for improved performance, data also suggest that air carriers may also need to switch service to higher-demand markets.

Among under-utilized airports, the data suggests that some have the conditions necessary for attracting increased air service and can potentially see improved levels of capacity utilization in the future. Others are already well-served and are less likely to see improved capacity utilization.

4.4 Case Study 3: Re-design of an Existing Benchmark

The analysis in sections 2 and 4.1 has in several steps shown the importance of determining the stakeholders and model in an airport benchmark:

- Section 2.2.2 shows that a variety of different performance metrics have been used in past studies with no motivation for why these metrics were selected and how they relate to the goals of the airport.

- Section 2.2.5 shows that past airport benchmarks have applied a variety of different DEA models, even though these studies have addressed the same general problem of comparing the efficiency of airports.

- Section 4.1 indicates that past airport benchmarks have been in alignment with the DEA framework and heuristics in some respects, but also that misalignment exists in several areas.

- (Schaar & Sherry 2008) showed that the choice of benchmark model can have a drastic impact on the study results, even to the degree of completely reversing the benchmark results.

In light of these facts, this case study contrasts the results from a past airport benchmark with the results of taking the premise of that original study, but applying the new airport benchmarking methodology in computing benchmark results.

The study reviewed in the analysis of the impact of benchmark model choices in (Schaar & Sherry 2008) was an analysis of the impact of airport size on efficiency (Bazargan & Vasigh 2003). In the present case study, that same analysis is used as the target for applying the new airport benchmarking methodology.

The objectives of this case study are to:

- Analyze how the parameters of the original analysis (performance metrics and benchmark model choice) differ from the parameters of the new study using the new benchmarking methodology.

- Study commonalities and differences between the results of the two studies.
- Identify which airports perform well and which ones do not using the new benchmark methodology.

The case study was conducted through the following steps:

1. Re-run the original case study on the OEP-35 airports. The reasons for re-running the analysis are that i) the original analysis was conducted on data spanning from 1996 to 2000 and newer data is now available; and ii) the original study did not address all of the OEP-35 airports and also included some non OEP-35 airports. For consistency with the other case studies in the dissertation, the new case study focused on performance in 2005-2008 for the OEP-35 airports.
2. Analyze the objectives that are related to the performance metrics in the original case study. Since the original case study did not discuss the objectives that its metrics address, this has to be "reverse-engineered". Once the objectives have been identified, a new benchmark can be designed that comprehensively addresses the relevant stakeholders' goals by using the new benchmark

methodology. The new benchmark methodology is also used to identify the benchmark model appropriate for the analysis.

3. Compute new benchmark results and i) investigate the underlying causes of the benchmark results and ii) compare them to the original benchmark.

In the remainder of the case study, the re-run of the original case study will be referred to as the “original study” and the benchmark using the new methodology will be referred to as the “new study”.

This section is organized as follows: The first subsection presents the original study. The second subsection applies the new benchmark methodology to design the new study. The third subsection presents the data and results of the new study and investigates the causes underlying the results. The fourth subsection compares the original and new case study. Finally, the last subsection presents conclusions.

4.4.1 Original Study

The original study did not provide a discussion of why the performance metrics used were chosen, nor did the study explain why the DEA model used was chosen. As a result, this section does not discuss the underlying reasons for the study design, but rather covers which performance parameters were used, how the benchmark results were computed, and shows the results.

4.4.1.1 Performance Metrics Used

This section presents the metrics used, the sources of data, and shows an overview of the performance data.

4.4.1.1.1 Metric Definitions

The original study used four inputs:

- **Runways**: The number of runways in use at the airport.
- **Gates**: The number of gates at the airport.
- **Operating costs**: Operating costs are on-going costs for operating the airport. They include eight sub-categories (Federal Aviation Administration 2001):
  - Personnel compensation and benefits
  - Communications and utilities
  - Supplies and materials
  - Repairs and maintenance
  - Contractual services
  - Insurance, claims and settlements

    - Miscellaneous
    - Other
- **Non-operating costs**: These are costs relating to the financing of the airport. They include (Federal Aviation Administration 2001):
    - Interest expense
    - Other

The study used six outputs:

- **Passenger volume**: This is the total number of enplaned passengers, encompassing both domestic and international passengers.
- **Air carrier aircraft movements**: This is the total number of takeoffs and landings by air carriers.
- **Other aircraft movements**: This is the total number of takeoffs and landings by general aviation, military, and other non-air carrier operators.
- **Portion of flights on-time**: This is the average portion of flights that depart from and arrive at the airport on-time. A flight is considered on-time if it is within 15 minutes of its scheduled time.

- **Aeronautical revenue**: The aeronautical revenues are receipts relating to aircraft operations at the airport. The revenues include (Federal Aviation Administration 2001):
    - Landing fees
    - Terminal/international arrival area rental or other charges
    - Apron charges/tiedowns
    - FBO revenue: contract or sponsor-operated
    - Cargo and hangar rentals
    - Aviation fuel tax retained for airport use
    - Fuel sales net profit/loss or fuel flowage fees
    - Miscellaneous
    - Other
- **Non-aeronautical revenue**: The non-aeronautical revenues come from a variety of sources that are not immediately related to aircraft operations. They include (Federal Aviation Administration 2001):
    - Land and non-terminal facilities
    - Terminal - food and beverage
    - Terminal - retail stores
    - Terminal - other

- Rental cars
- Parking
- Miscellaneous
- Other

4.4.1.1.2 Data Sources

The data for the ten parameters was compiled from several different sources:

Table 4.21 - Data sources for original case study

Metric	Source
Runways	The FAA's National Plan of Integrated Airport Systems (FAA 2008)
Gates	This was compiled through a large number of sources, including (A-Z World Airports 2010) and airport websites.
Operating costs	The FAA's Compliance Activity Tracking System, to which all airports must report their financial performance (Federal Aviation Administration 2010a)
Non-operating costs	The FAA's Compliance Activity Tracking System (Federal Aviation Administration 2010a)
Passenger volume	The T100 database, which is compiled from data collected by Office of Airline Information

Metric	Source
	(OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b)
Air carrier aircraft movements	The FAA's Air Traffic Activity System (Federal Aviation Administration 2010)
Other aircraft movements	The FAA's Air Traffic Activity System (Federal Aviation Administration 2010)
Portion of flights on-time	The on-time database, compiled from data collected by the OAI at the BTS (Bureau of Transportation Statistics 2010b). This data only encompasses U.S. carriers.
Aeronautical revenue	The FAA's Compliance Activity Tracking System (Federal Aviation Administration 2010a)
Non-aeronautical revenue	The FAA's Compliance Activity Tracking System (Federal Aviation Administration 2010a)

The purpose of this re-run of the original study is to provide a point of comparison to the new case study. Accordingly, the inputs and outputs are treated in the same way as they were in the original study. A discussion of some of the issues of this performance data is provided in section 4.4.4.

4.4.1.1.3 Summary of Performance Data

Although the analysis was conducted on an annual basis, this section presents average values for each of the inputs and outputs across the whole period 2005-2008. The full details of the inputs and outputs are provided in Appendix E.

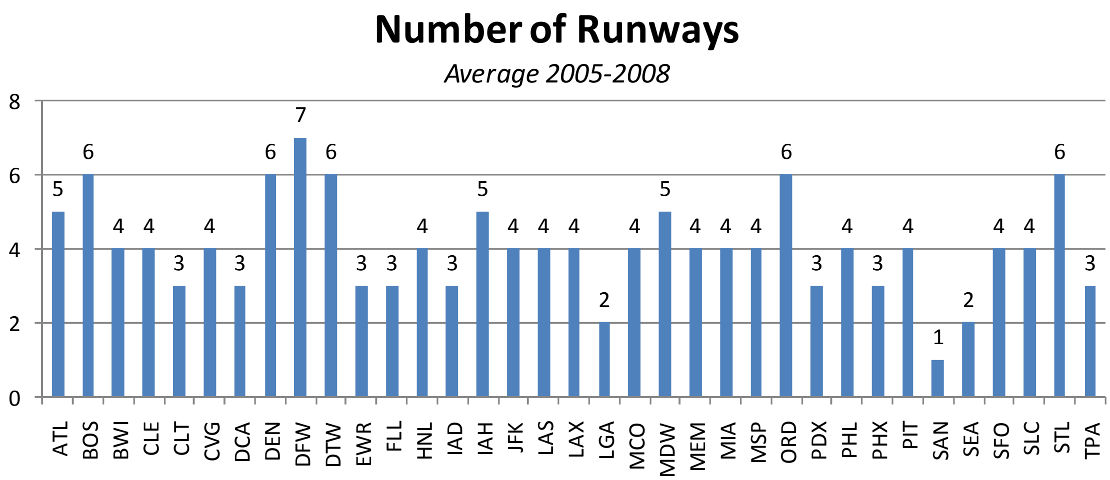

Figure 4.31 - Number of runways

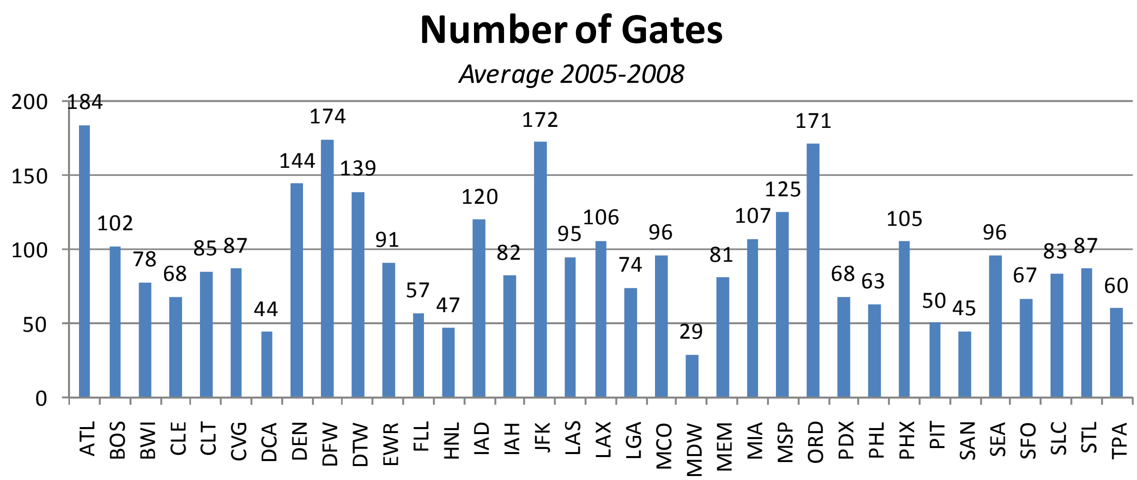

Figure 4.32 - Number of gates

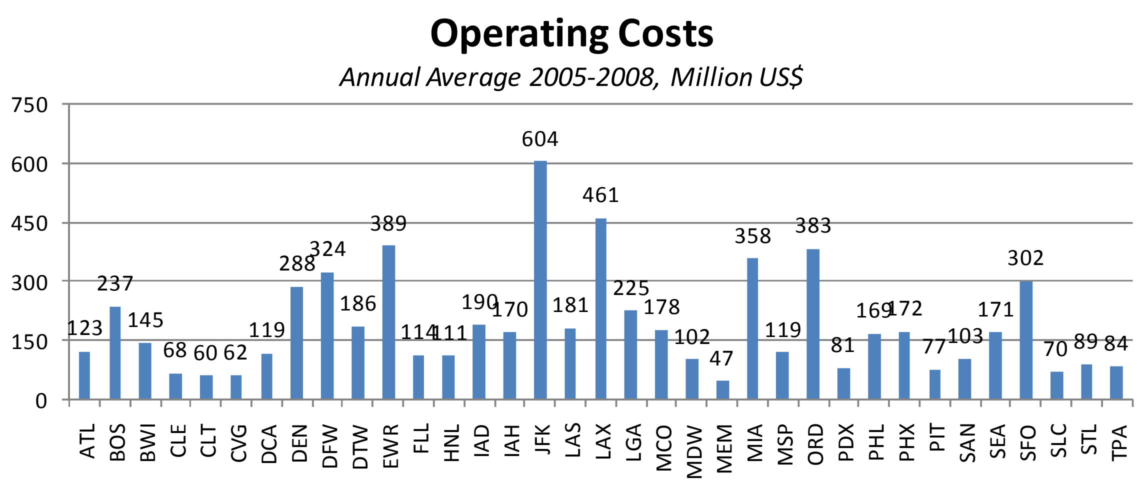

Figure 4.33 - Annual operating cost, average 2005-2008, million US$

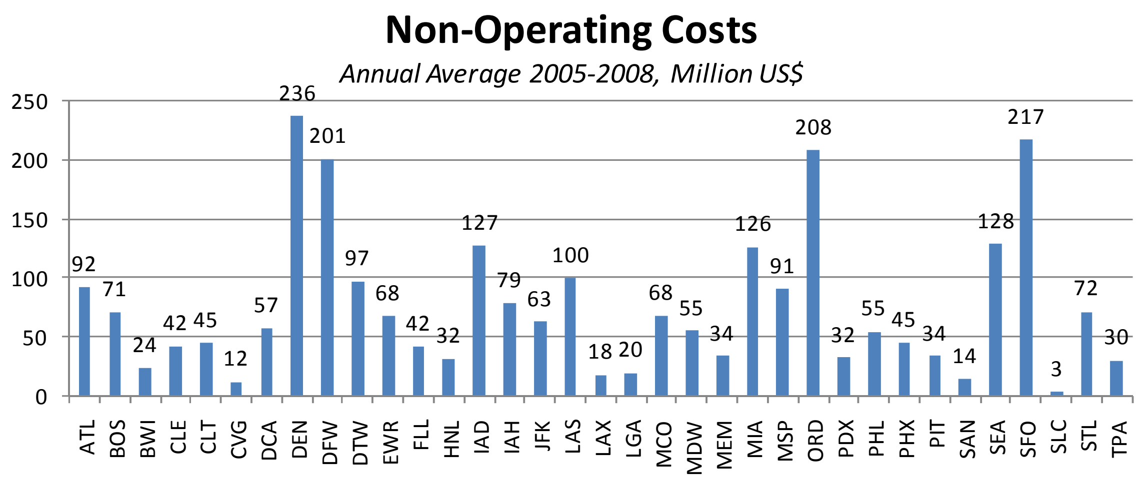

Figure 4.34 - Annual non-operating cost, average 2005-2008, million US$

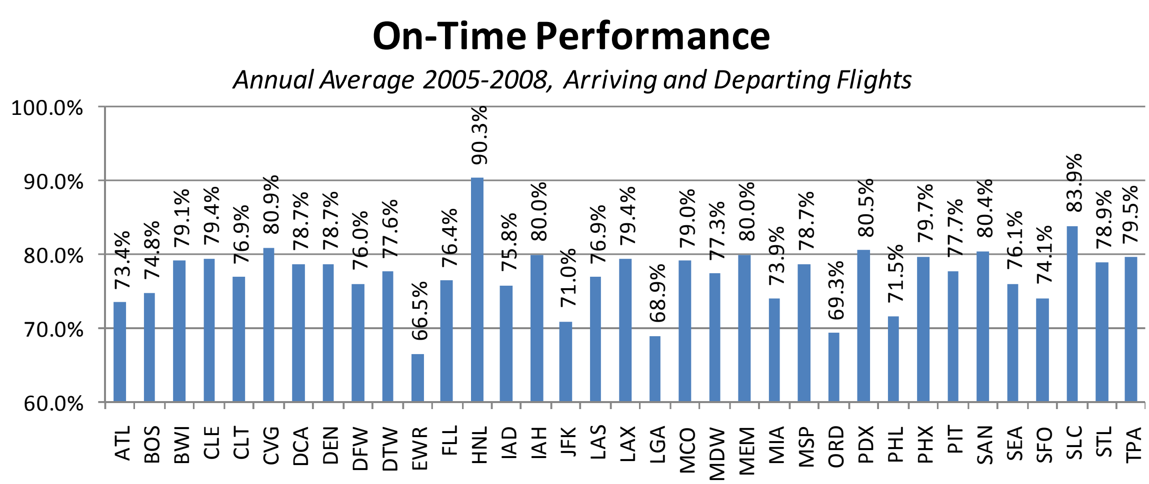

Figure 4.35 - Portion of flights arriving and departing on-time, average 2005-2008

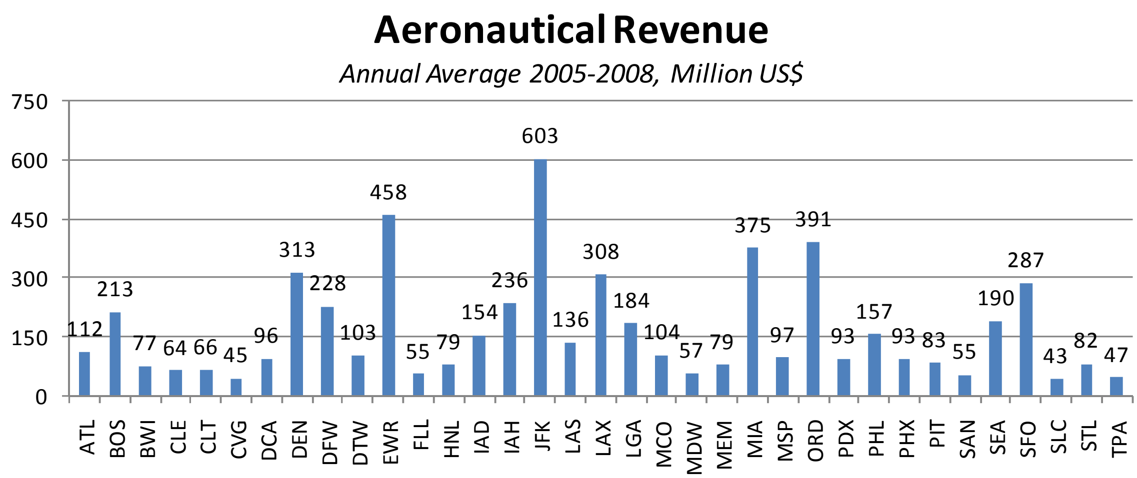

Figure 4.36 - Annual aeronautical revenue, average 2005-2008, million US$

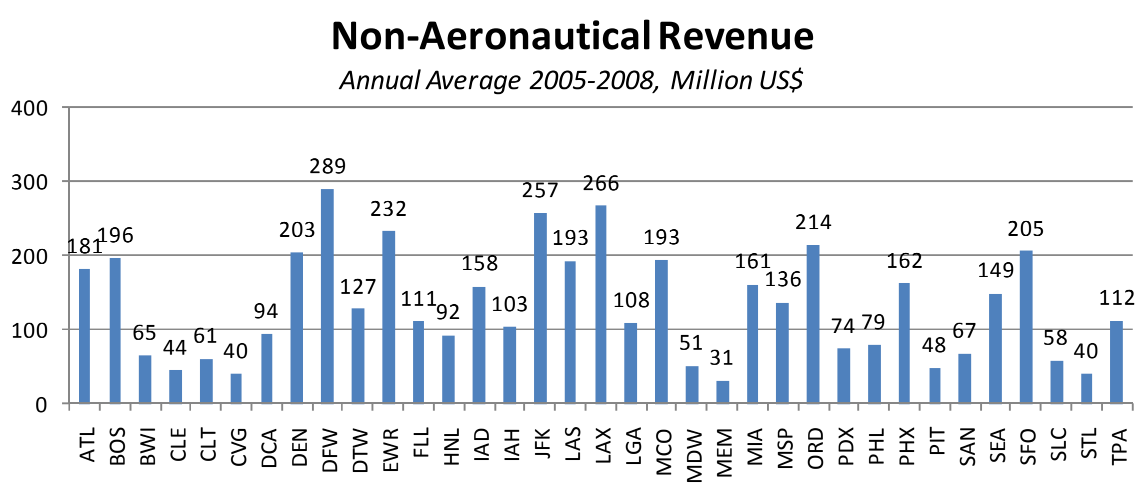

Figure 4.37 - Annual non-aeronautical revenue, average 2005-2008, million US$

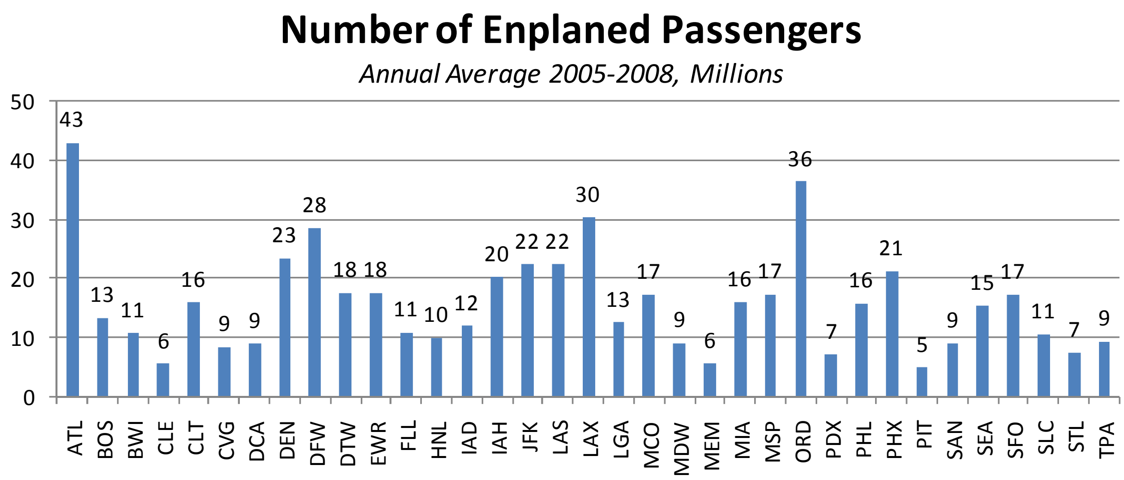

Figure 4.38 - Annual number of enplaned passengers (international and domestic), average 2005-2008, millions

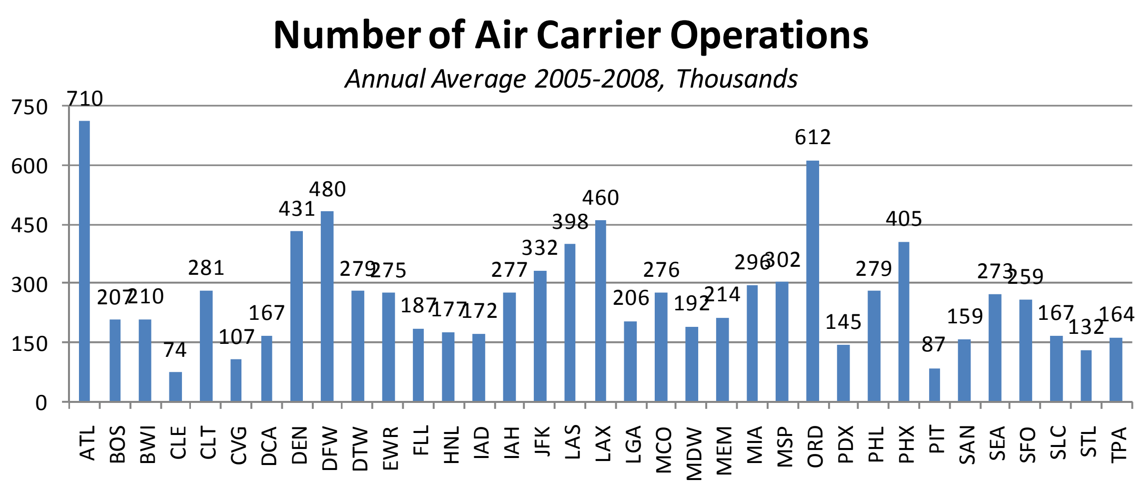

Figure 4.39 - Annual number of air carrier operations (takeoffs and landings), average 2005-2008, thousands

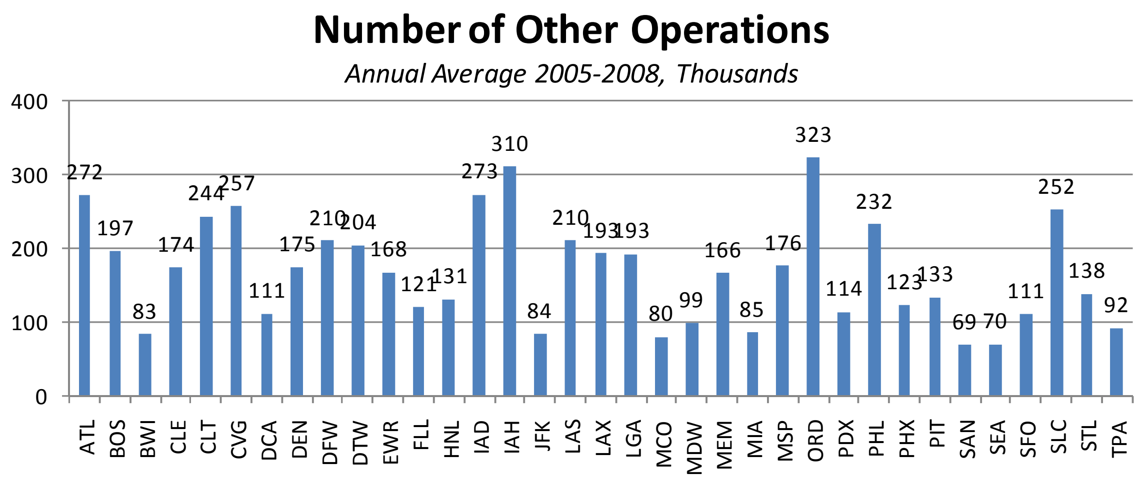

Figure 4.40 - Annual number of other operations (including general aviation and military flights; takeoffs and landings), average 2005-2008, thousands

4.4.1.2 Benchmark Model Used

Although the original study did not systematically address the selection of benchmark model, the original study's model choice can be characterized using the DEA model framework from Appendix A. The model choices are presented in Table 4.22.

Table 4.22 - DEA model choices in original study

Scalarizing function			Technology			Timespan	Tie-breaking
Aggre-gation	Weights	Orient-ation	Returns to scale	FDH	Integrality		
Maximin	Specific weights	Input oriented	CRS	No use of FDH	No integrality constraints	No use of Malmquist index; simply one analysis per year	Super-efficiency using artificial DMU

4.4.1.3 Benchmark Results

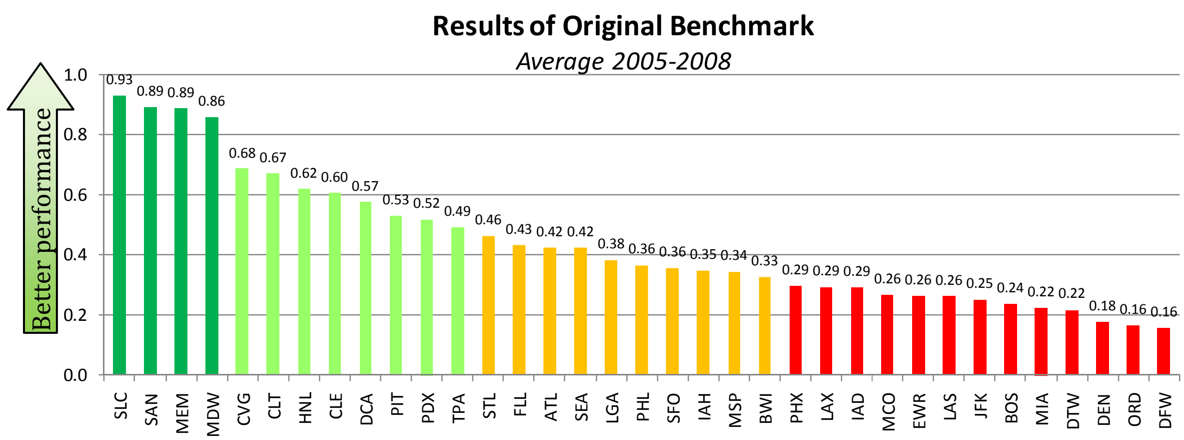

Figure 4.41 - Results of original benchmark, average 2005-2008

The benchmark results are shown in Figure 4.41, with the best level of performance for SLC, SAN, and MEM, while the worst performance is recorded for DFW, ORD, and DEN. The full details of the results are provided in Appendix E.

The original study found that for the period 1996-2000 and for the airports reviewed, small airports had greater levels of efficiency than large airports. To test whether that conclusion holds in the re-run using new data and a somewhat different set of airports, a test was conducted using the number of enplaned passengers as an indicator of airport size. The data was plotted as shown in Figure 4.42 and then two groups of airports (large and small) was created and a Kruskal-Wallis test was conducted on the efficiency ranks.

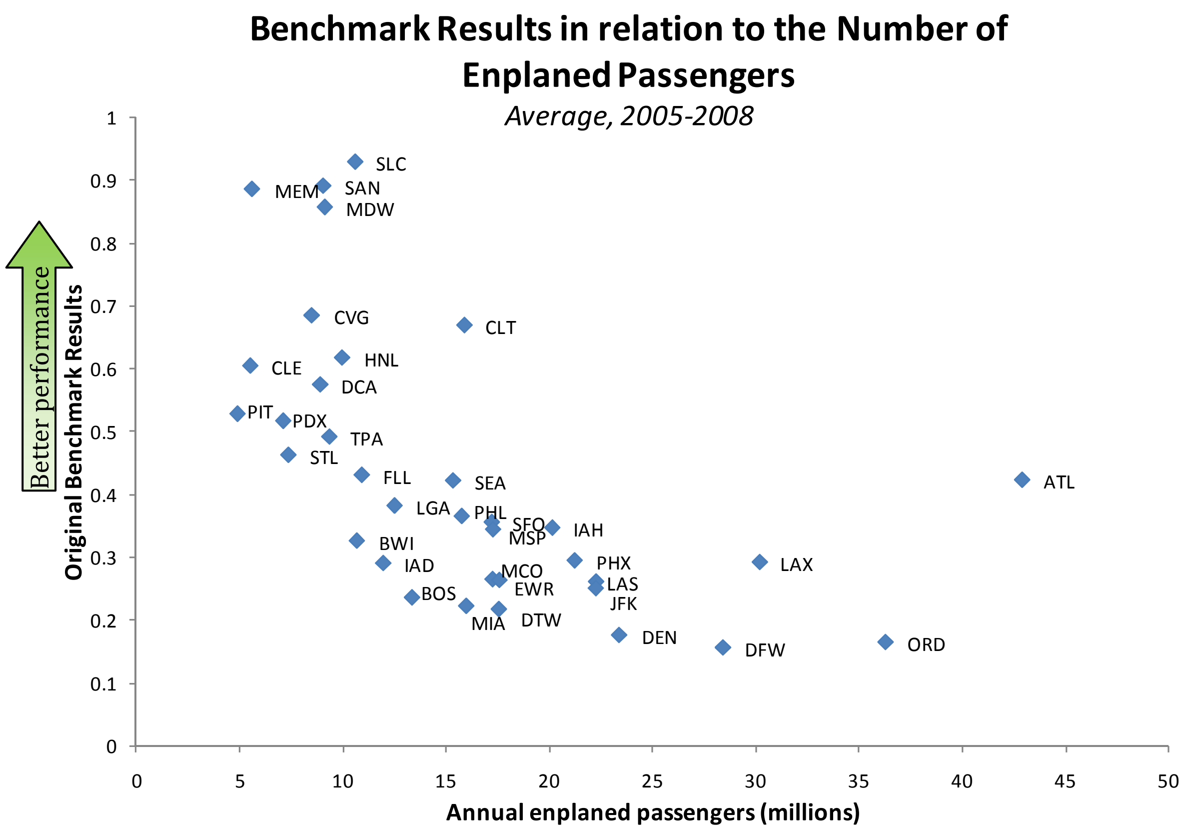

Figure 4.42 - Benchmark results as a function of the number of enplaned passengers

To test the difference in efficiency between large and small airports, the airports were divided into two groups of nearly equal size based on the number of enplaned passengers (because the total number of airports was 35, the first 17 airports were classified as small and the remaining 18 were classified as large). It is not possible to assume that the efficiency scores are normally distributed which rules out the t-test for comparing the performance of the two groups. Instead the Kruskal-Wallis test is performed on the ranks of the efficiency score. The Kruskal-Wallis test does not require that the population have a normal distribution. The mean ranks, χ^2, and significance values from the test are presented in Table 4.23.

Table 4.23- Kruskal-Wallis test on benchmark result rankings

Year	Mean rank		Chi-square	Asymptotic significance
	Small	Large		
2005	11.18	24.44	14.658	0.00013
2006	11.71	23.94	12.472	0.00041
2007	11.29	24.33	14.157	0.00017
2008	11.35	24.28	13.910	0.00019

The results show that the difference is significant at the 95% level for each year in the study, with small airports showing better performance than large airports. These results support the original study's finding for 1996-2000 that larger airports exhibit lower levels of efficiency.

Beyond these findings, it is not the purpose of this case study to analyze the findings and underlying reasons for the performance outcomes of the original benchmark. However, a contrast between these original benchmark results and the results of the new study is presented in section 4.4.4.

4.4.2 Design of New Study

The process for designing the new case study is described in Figure 4.43 and is detailed in the following subsections.

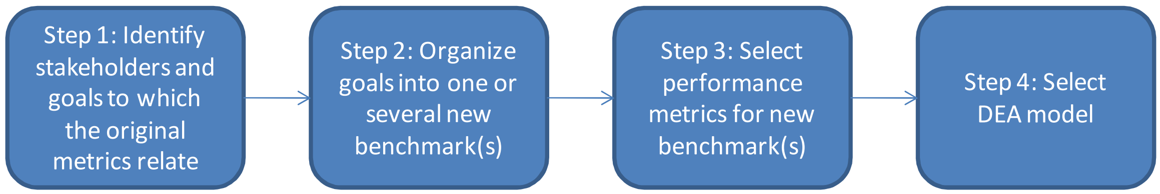

Figure 4.43- Process for designing new study

4.4.2.1 Stakeholders and Goals Related to Original Study Metrics

The original study did not explain why the performance metrics used were selected. Understanding to which stakeholders these metrics are relevant, and to which goals they relate, requires an amount of "reverse engineering". In this section, the metrics are categorized and analyzed to determine a mapping to stakeholders and goals. Table 4.24 shows the categorization of the original study's performance metrics.

Table 4.24 - Categorization of performance metrics in original study

Metric	Category
Runways	Infrastructure capacity
Gates	Infrastructure capacity
Operating costs	Financial performance
Non-operating costs	Financial performance
Passenger volume	Throughput
Air carrier aircraft movements	Throughput
Other aircraft movements	Throughput
Portion of flights on-time	Congestion costs
Aeronautical revenue	Financial performance
Non-aeronautical revenue	Financial performance

This categorization shows two separate themes among the metrics:

1. **Operational efficiency**: Infrastructure capacity, throughput, and congestion costs are metrics that relate to the operational efficiency of the airport. The operational efficiency of the airport can be described as the volume of throughput that can be achieved using the available

infrastructure capacity while minimizing the undesirable “side-effect” of congestion costs.

2. **Investment quality**: The investment quality of the airport – i.e. the level of attractiveness of the airport to investors – is determined by considering factors such as the airport’s financial performance and the characteristics of its volumes of throughput.

Referring to the model of airport stakeholders and their goals for the airport as described in section 2.1.3, these two themes can be tied to the stakeholders and goals shown in Table 4.25.

Table 4.25 - Mapping of goals to stakeholder for new case study

Overall goal	Stakeholder	Relevant goals of stakeholder (ref section 2.1.3)	Commentary on why this stakeholder goal is relevant
Maximize operational efficiency while minimizing delay effects	Airport organization	- Ensure sufficient (but not excessive) infrastructure capacity	While the airport organization wants to ensure the capacity is used to as high a degree as possible, it should not be over-utilized to the degree that severe congestion occurs.
	Air carriers	- Ensure on-time performance - Ensure low cost of operations	Air carriers want operational efficiency in the form of low delays but also low costs of operations, which result from the efficient use of airport infrastructure since high capacity utilization results in a low cost per use.
	Federal government	- Ensure safety, security, and efficiency of operations	Operational efficiency is one of the federal government's directly stated goals.
Maximize investment quality	Investors and bond-holders	- Optimize performance on factors affecting credit ratings	Maximizing investment quality directly maps to the interests of investors and bond-holders
	Airport organization	- Ensure sufficient (but not excessive) infrastructure capacity - Manage costs	Achieving good performance in terms of investment quality allows the airport organization to build more capacity when needed thanks to the availability of capital, and having a good credit rating results in lower costs through better interest rates.

4.4.2.2 Organizing Goals into New Benchmarks

The two, distinct goals described in Table 4.25 have different stakeholders. To clearly gauge each airport's ability to accomplish each of the two goals, they will be treated as two separate benchmarks in the remainder of the case study. They will be referred to as the "operational efficiency benchmark," and the "investment quality benchmark," respectively.

4.4.2.3 Determine Performance Metrics for New Study

This section describes the process for and results of determining the performance metrics in the two component benchmarks.

4.4.2.3.1 Operational Efficiency Benchmark

The goal of the operational efficiency benchmark is to "maximize operational efficiency while minimizing delay effects". The operational efficiency in the context of the original benchmark pertains to maximizing throughput given the available capacity.

The throughput measures for the airport are defined in terms of passengers and in terms of aircraft movements, as illustrated in the original benchmark.

Although a correlation between the two output metrics exists since more aircraft can carry more passengers, using only one of the two metrics may ignore some stakeholders' interests. A focus exclusively on passengers would ignore the importance of the number of aircraft movements per unit of capacity to the level of capacity utilization, while a focus only on the number of aircraft movements fails to address the ultimate goal for some stakeholders of moving as many passengers as possible. As a result, the study includes both the throughput measures in the benchmark. Specifically, the study uses:

- The number of enplaned domestic and international passengers, including both O&D and connecting passengers.
- The number of domestic and international operations, including air carrier operations as well as general aviation and other types of operations.

As discussed further in section 4.3.2.2.3, the primary factor determining the amount of traffic that an airport can handle is the runway capacity. This is the main throughput bottleneck in the airport system (Neufville & Odoni 2003, p. 367).

Because of the presence of several factors impacting each runway's capacity as described in section 4.3.2.2.3, the number of runways alone does not determine the airport's capacity, as was done in the original study. Instead, a measure of the actual capacity of the set of runways as proposed in (Kumar & Sherry 2009) is used in this case study. In this measure, airport Capacity Coverage Charts (CCCs) are used along with data on the costs of delay to determine average airport capacity. CCCs describe how much runway capacity was available and for how long (Neufville & Odoni 2003, p. 402). This average airport capacity is the measure used in this analysis.

The cost of delay can be measured either in terms of aircraft-based delays or in terms of passenger-based delays. Aircraft-based delays include five different metrics (Federal Aviation Administration 2010a):

- **Airport Departure Delay:** "The actual wheels off minus the scheduled gate out plus the unimpeded taxi out time, in minutes. Negative values contribute to the total."
- **Gate Departure Delay:** "The sum of minutes of gate departure delay of 1 minute or more departures. Gate departure delay is the difference between the actual gate out time and scheduled or flight plan gate out time, in minutes."

- **Taxi Out Delay:** "The sum of minutes of taxi out delay of 1 minute or more. Taxi out delay equals actual taxi out time minus unimpeded taxi out time."
- **Taxi In Delay:** "The sum of minutes of taxi in delay of 1 minute or more. Taxi in delay equals actual taxi in time minus unimpeded taxi in time."
- **Gate Arrival Delay:** "The sum of minutes of gate arrival delay of 1 minute or more. Gate arrival delay is the difference between the actual gate in time and the scheduled or flight plan gate in time."

Generalized passenger-based delay metrics were introduced in (Wang & Sherry 2007). Although passenger delay metrics could include both arrival and departure delay metrics, the primary concern of most passengers is arriving on time (and some departure delays can be compensated for by "making up" time during the enroute portion of the flight). Accordingly, the primary focus of passenger delay metrics is arrival delay.

In the choice between the different categories of metrics, the final selection was the passenger-based arrival delay metric. The motivation for this choice was the view of the passenger as the end consumer of the air service, rather than the air carrier.

To test the impact of the delay metric choice, a correlation test was done for all five parameters under consideration. The Pearson correlation coefficients for these tests are presented in Table 4.26.

Table 4.26 - Pearson correlation coefficients for test of correlation between delay metrics. The delay data was computed annually for 2005-2008 for the OEP-35 airports, and the total delay for all flights was summed for each year.

	Passenger arrival delay	Passenger departure delay	Gate departure delay	Taxi out delay	Airport departure delay	Taxi in delay	Gate arrival delay
Passenger arrival delay							
Passenger departure delay	r = 0.99 p = 2.2e-16						
Gate departure delay	r = 0.87 p = 2.2e-16	r = 0.88 p = 2.2e-16					
Taxi out delay	r = 0.78 p = 2.2e-16	r = 0.78 p = 2.2e-16	r = 0.89 p = 2.2e-16				
Airport departure delay	r = 0.85 p = 2.2e-16	r = 0.86 p = 2.2e-16	r = 0.97 p = 2.2e-16	r = 0.96 p = 2.2e-16			
Taxi in delay	r = 0.83 p = 2.2e-16	r = 0.84 p = 2.2e-16	r = 0.91 p = 2.2e-16	r = 0.80 p = 2.2e-16	r = 0.89 p = 2.2e-16		
Gate arrival delay	r = 0.91 p = 2.2e-16	r = 0.90 p = 2.2e-16	r = 0.95 p = 2.2e-16	r = 0.91 p = 2.2e-16	r = 0.96 p = 2.2e-16	r = 0.84 p = 2.2e-16	

The correlation coefficients show that all delay metrics are strongly correlated (0.78 or higher) and that many are very highly correlated (0.90 or higher). All are significant at the 95% confidence level. This suggests that in the event that another delay metric had been chosen, the benchmark results would have been similar.

In summary, this benchmark uses the conceptual ratio of (enplaned passengers, aircraft movements) : (capacity, passenger arrival delay).

4.4.2.3.2 Investment Quality Benchmark

The level of investment quality of airports is gauged by the credit rating agencies, as described in section 2.1.3.2.6. (Forsgren 2007) provides details on the factors that are considered in the credit rating process. These factors are listed in Table 4.27 with categorization added by this author.

Table 4.27 - Factors considered in airport credit ratings (Forsgren 2007). Categorizations added by this author.

Category	Factor
Regional growth	Historical/ projected population growth
	Historical/ projected employment growth
Air service	Historical/ projected passenger growth
	Portion of traffic that is O&D
	The role of the airport in the dominant carrier's network
	The importance of the airport to the overall NAS
	Financial strength of carriers with large amounts of connecting traffic
	Is the airport in a favorable geographic location (i.e. natural hub)?
Airfield/facilities	Capacity utilization
	Attractiveness of facilities
Financial factors	Debt burden and carrying costs
	Non-aeronautical revenues
	Ability to change rates

Table 4.27 lists 13 factors that are considered by credit rating agencies in determining the investment quality of an airport. Since only 35 airports are included in the benchmark, a single benchmark considering all of these factors simultaneously is not feasible due to the low resulting ratio of DMUs to parameters in the benchmark (a commonly applied rule of thumb is that the number of DMUs should be at least twice the product of the number of inputs and outputs (R. G. Dyson et al. 2001)). Instead, each category of parameters will be treated in a separate component benchmark to determine the relative performance of each airport.

To conduct the benchmark, the factors listed in Table 4.27 had to be translated into performance metrics which could be included in each benchmark. Table 4.28, Table 4.29, Table 4.30, and Table 4.31 present the results of this translation. Some of the factors considered by the credit rating agencies were qualitative in nature and some were not possible to translate into comparative performance metrics using available data. Where that was the case it has been noted in the tables. As Table 4.30 shows, this resulted in a benchmark of the Airfield/facilities category not being possible to complete.

Table 4.28 - Regional growth factors

Factor	Metric	Discussion
Historical/ projected population growth	Population growth for the airport's Metropolitan Statistical Area (MSA).	This is a direct indicator of the region's population growth.
Historical/ projected employment growth	Growth of the MSA's regional GDP.	The underlying assumption is that growth in the region's economy is correlated with employment growth.

Table 4.29 - Air service factors

Factor	Metric	Discussion
Historical/ projected passenger growth	Growth of enplaned passengers.	This is a direct indicator of passenger growth.
Portion of traffic that is O&D	O&D passengers as percentage of all enplaned passengers.	This is a direct indicator of the portion of O&D traffic.

Factor	Metric	Discussion
The role of the airport in the dominant carrier's network	Total number of enplanements at this airport for the carrier as percentage of carrier's total enplanements, for the carrier with the largest number of enplanements at this airport.	The underlying assumption with this metric is that a hub which represents a large portion of a carrier's traffic is more important to the carrier and thereby more stable and is less likely to be subject to cuts than a hub with secondary status for the carrier.
The importance of the airport to the overall NAS	This airport's percentage of all enplaned passengers at OEP-35 airports.	The underlying assumption is that airports with a high portion of passengers have a high importance to the overall NAS.
Financial strength of carriers with large amounts of connecting traffic	N/A	This is a factor that requires deep financial analysis of individual air carriers and it is outside the scope of this analysis.
Is the airport in a favorable geographic location (i.e. natural hub)?	N/A	This is a factor that requires expert judgment to determine a comparative numerical value. It is outside the scope of this analysis.

Table 4.30 - Airfield/facilities factors

Factor	Metric	Discussion
Capacity utilization	N/A	This factor is addressed in the capacity utilization and operating efficiency benchmarks described in sections 4.3 and 4.4.2.3.1.
Attractiveness of facilities	N/A	This is a qualitative factor that is outside the scope of this analysis.

Table 4.31 - Financial factors

Factor	Metric	Discussion
Debt burden and carrying costs	Debt service coverage ratio	This is a metric that indicates the ratio between the airport's operating surplus and the cost of servicing its debt. The formula for computing this metric is described in Figure 4.44.
Non-aeronautical revenues	Non-aeronautical revenues as a percentage of total revenues.	This is a direct indicator of the portion of non-aeronautical revenues.
Ability to change rates	N/A	This is a qualitative factor that requires expertise in each airport's legal environment to determine. It is outside the scope of this analysis.

$$\text{Debt service coverage ratio} = \frac{((\text{Total Operating Revenues} + \text{Interest Income}) + (\text{Annual PFC Revenue}) - (\text{Operating and Maintenance Expenses} - \text{Depreciation and amortization}))}{\text{Gross Annual Debt Service Requirement}}$$

Figure 4.44 - Debt service coverage ratio calculation (Pezzimenti & Macdonald 2005)

In summary, the investment quality benchmark is made up of three component benchmarks, and each of these benchmarks is comprised only of output metrics which should be maximized. Furthermore, all of the metrics used in the benchmarks are percentages or other ratios that are already scale-independent (meaning that they do not scale up/down with larger/smaller-sized airports). Accordingly, in the DEA implementation, this is treated as all airports having the same, single, constant input which can be assigned any positive value. In this implementation, that constant is assigned the value 1. The three benchmarks are:

1. **Regional growth benchmark**: The regional growth benchmark can be described by the conceptual ratio (population growth; regional GDP growth) : (1).

2. **Air service benchmark**: The air service benchmark can be described by the conceptual ratio (passenger growth; O&D passenger percentage; portion of main carrier's passengers enplaned; portion of OEP-35 passengers) : (1).
3. **Financial factors**: The financial factors benchmark can be described by the conceptual ratio (debt service coverage ratio; non-aeronautical revenue percentage) : (1).

As the preceding tables show, these three benchmarks are not complete in covering all the factors considered by credit rating agencies due to limitations on data availability and expert judgment. However, they do cover a majority (eight of 13) of the factors considered by the credit rating agencies. It should also be noted that although these are the 13 factors explicitly listed, there is no limitation that keeps credit rating agencies from considering further factors beyond these 13.

4.4.2.4 Select DEA Model for New Study

This section describes the model selection for each of the new studies.

4.4.2.4.1 Operational Efficiency Benchmark

This benchmark gauges the conceptual ratio of (enplaned passengers, aircraft movements) : (capacity, passenger arrival delay). The units of these metrics are passengers, aircraft, aircraft movements, and hours of delay, respectively. As with the previous case studies, this results in DEA being the appropriate modeling choice for the benchmark.

The results of the application of the framework and heuristics to determine a specific DEA model for this analysis are now presented.

- **Aggregation**: The heuristics prescribe an ε-maximin aggregation function as the default choice unless negative data is present or if any reasons exist why slacks cannot be ignored. Neither of those conditions are met in this analysis, and accordingly an ε-maximin aggregation is used.
- **Weights**: The heuristics prescribe the use of specific weights unless any reasons are present for choosing range-adjusted weights. The specific weights allow each DMU to select its own optimal weights and in this study that is an appropriate selection to reflect the decisions of those involved in managing services at the airport. Specific weights

are used since no reasons for using range-adjusted weights are present.

- **Orientation**: The heuristics prescribe using an orientation based on the factors which are most controllable by airport management. In this analysis, the outputs are passenger and aircraft movements, which cannot be directly controlled by airport management but can be indirectly influenced through marketing and incentive campaigns. The inputs are deemed to be less within the control of airport management: Runway capacity is largely a static value which is difficult to influence; once a runway has been constructed it is difficult to remove it (Martín & Román 2001, pp. 152-153), and conversely at some airports space constraints and community opposition limit the ability to add further runway capacity (Neufville & Odoni 2003, p. 168). Delay levels can generally only be controlled by airlines through additions/reductions in traffic or by the U.S. Congress through the "High Density Rule", which caps the number of operations at certain airports (Neufville & Odoni 2003, pp. 474-475).
- **Returns to scale**: The heuristics prescribe that if modeling some version of labor and capital resources as inputs and passengers and

aircraft movements as outputs, then VRS should be used. The runway capacity can be considered a capital resource, and passengers and aircraft movements are used as an output, and accordingly VRS is the choice for this model.

- **FDH**: The Free Disposal Hull should be applied only if some reason exists why comparison only to observed combinations of inputs and outputs should be made, but no such reason exists in this analysis.
- **Integrality**: Integrality constraints should be applied in cases where input or outputs are indivisible into fractions and of low magnitude, and if large errors in the results would be introduced if these inputs or outputs were assumed to have decimal values. Both of the output parameters can only take on integer values since the number of passengers and aircraft movements are indivisible, but both of these have a very high magnitude and as a result, no integer constraints are necessary. On the input side, the number of delay hours can take on any continuous value, but the number of hourly aircraft movements can only take on integer values and has a low magnitude with a median value of 30, which causes this input to have be integer constrained.

- **Timespan**: If any key technology changes have occurred during the timespan being studied that would impact the ability of DMUs to achieve strong performance, then a Malmquist index method should be used. If not, the performance for each year can be analyzed independently. In this analysis, technology changes would imply some new technology being introduced which would allow a higher number of aircraft movements or passengers being moved using existing runway capacity while not increasing delays. Since the minimum aircraft separation standards remained constant during the 2005-2008 period, no increase in the number of aircraft per unit of runway capacity was possible. The number of passengers carried could be increased if a new aircraft able to carry more passengers than any other were introduced during the period of the analysis. The Airbus A380 is such an aircraft but it was in very limited service to U.S. airports during this period, beginning service only to JFK on August 1, 2008 (J. Lee 2008), causing its impact on performance during this period to be very limited. As a result of no major changes occurring in this time period, no Malmquist index calculation is necessary.

- **Tie-breaking**: The heuristics prescribe that tie-breaking be used only if a reason exists that all airports must be fully ranked. No such reason exists in this analysis.

Table 4.32 summarizes the modeling choices for this benchmark.

Table 4.32 - Modeling choices for benchmark

Scalarizing function			Technology			Timespan	Tie-breaking
Aggre-gation	Weights	Orient-ation	Returns to scale	FDH	Integrality		
ε-maximin	Specific weights	Output oriented	VRS	No use of FDH	Runway capacity integer constrained	No use of Malmquist index; simply one analysis per year	None

4.4.2.4.2 Investment Quality Benchmark

Each component of the investment quality benchmark is treated separately in the model selection stage. All three benchmarks use a combination of metrics of different units without a known utility function, and accordingly DEA models are used in each of the cases. The modeling parameters and data characteristics share sufficient similarity throughout the three component benchmarks that the model choices are the same in each benchmark, as described below:

- **Aggregation**: The benchmark data contains some parameters which take on negative values. The heuristics prescribe the use of the additive aggregation function with tolerance for negative values in this circumstance.
- **Weights**: The heuristics prescribe the use of specific weights unless any reasons are present for choosing range-adjusted weights. The specific weights allow each DMU to select its own optimal weights and in this study that is an appropriate selection to reflect the decisions of those involved in managing services at the airport. Specific weights are used since no reasons for using range-adjusted weights are present.
- **Orientation**: In each of the benchmarks, all the factors are outputs and each benchmark uses a single, constant input. As a result of this, all of the benchmarks are designed as output-oriented.
- **Returns to scale**: Because the benchmarks use ratios and percentages which have already been normalized for scale, no other economies of scale are present. Accordingly, the models are set to CRS.

- **FDH**: The Free Disposal Hull should be applied only if some reason exists why comparison only to observed combinations of inputs and outputs should be made, but no such reason exists in this analysis.
- **Integrality**: All of the benchmark parameters are percentages and ratios, all of which take on continuous values. No benchmark parameters have integer constraints.
- **Timespan**: As is discussed in section 4.4.3.2, the investment quality benchmark is conducted only for one single year. As a result, no Malmquist index calculation considerations apply.
- **Tie-breaking**: The heuristics prescribe that tie-breaking be used only if a reason exists that all airports must be fully ranked. No such reason exists in this analysis since the same credit ratings can be shared by multiple airports.

The modeling choices for each of the three component benchmarks in this case study are presented Table 4.33.

Table 4.33 - Modeling choices for benchmark

Scalarizing function			Technology			Timespan	Tie-breaking
Aggre-gation	Weights	Orient-ation	Returns to scale	FDH	Integrality		
Additive aggregation function with tolerance for negative values	Specific weights	Output oriented	CRS	No use of FDH	No parameters integer constrained	N/A – single time period	None

4.4.3 Data, Results, and Discussion of New Study

This section presents the data sources and a summary of the benchmark data used, benchmark results, and a discussion of the benchmark results. The operational efficiency benchmark is treated first and the investment quality benchmark second.

4.4.3.1 Operational Efficiency Benchmark

The operational efficiency benchmark is computed for each year 2005-2008 to match the scope of each of the previous benchmarks.

4.4.3.1.1 Data Sources

Several data sources were used for this benchmark:

- **Enplaned passengers**: This data was derived from the T100 database, which is compiled from data collected by the Office of Airline Information (OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b)
- **Aircraft movements**: This data was compiled from the FAA's Air Traffic Activity System (Federal Aviation Administration 2010).
- **Runway capacity**: As in section 4.3.2.4.1, this data was derived from the analysis described in (Kumar & Sherry 2009). This analysis in turn was conducted using the Aviation System Performance Metrics (ASPM) database (Federal Aviation Administration 2010d) along with the T100 database and the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c).
- **Passenger arrival delay**: The passenger delay data was retrieved from the on-time database, compiled from data collected by the OAI at the BTS (Bureau of Transportation Statistics 2010b), and from the T100 database. This data only encompasses U.S. carriers. The delay was computed as the difference between each flight's scheduled arrival time with its actual arrival time for those flights delayed more

than 15 minutes, and then multiply that difference by the number of passengers on each flight (Wang & Sherry 2007, p. 3).

4.4.3.1.2 Summary of Benchmark Parameters

This section presents an average for each of the four parameters used in the benchmark for the period 2005-2008. The full details of the inputs and outputs are provided in Appendix E.

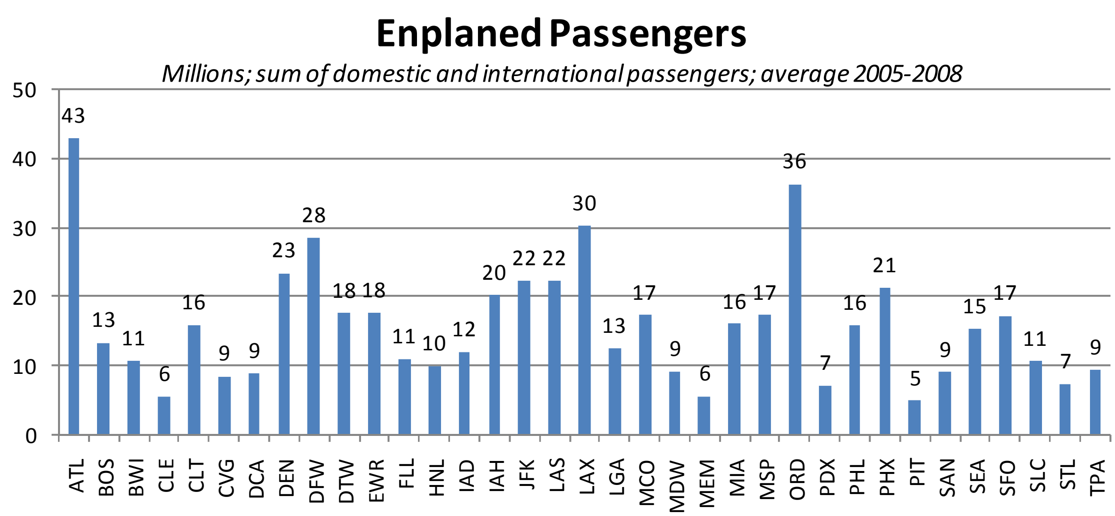

Figure 4.45 - Total number of enplaned passengers (millions). Including both domestic and international passengers, average 2005-2008.

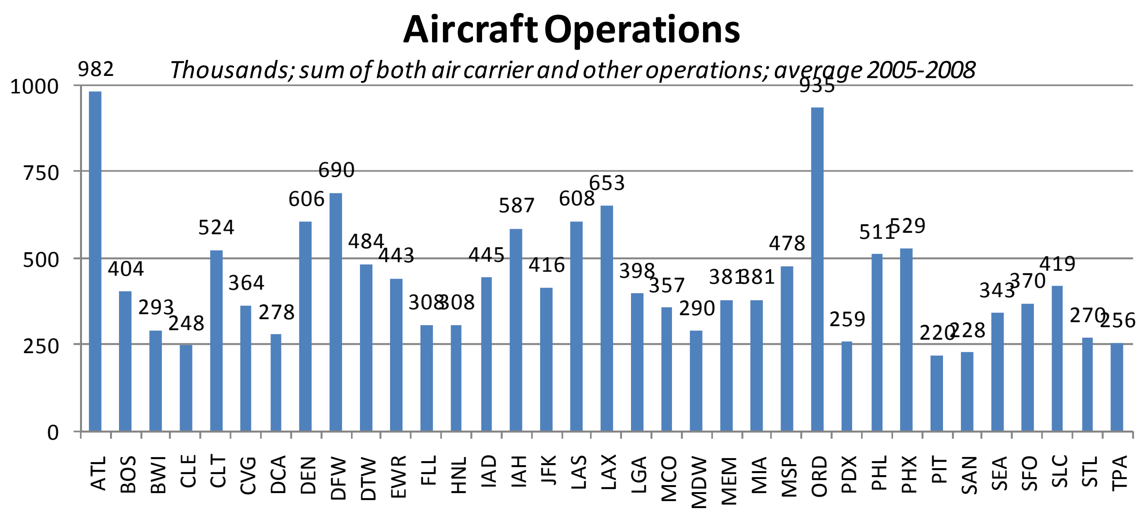

Figure 4.46 - Total number of aircraft operations, including both takeoffs and landings, and including both air carrier and other operations. Average for 2005-2008.

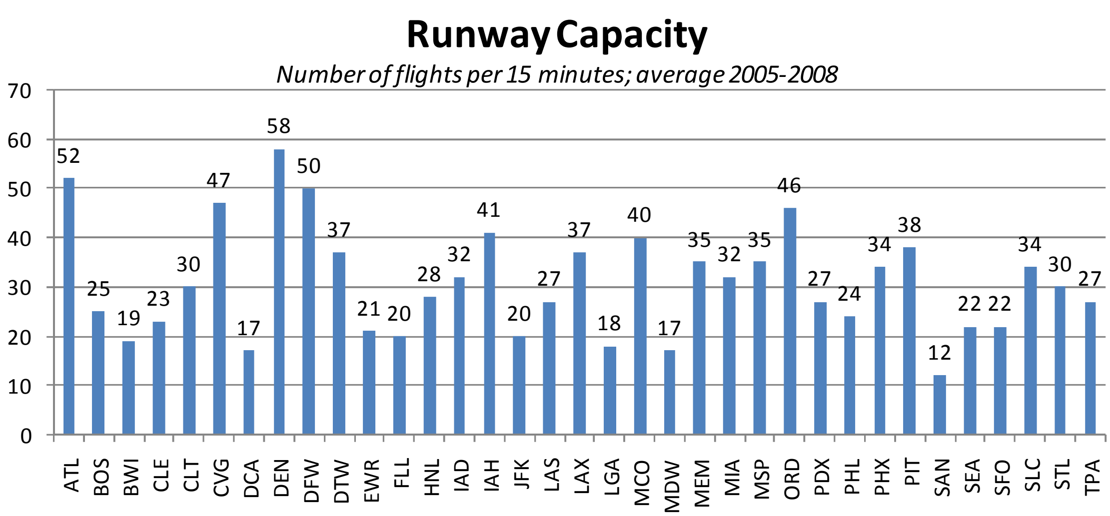

Figure 4.47 - Runway capacity. Number of flights per 15 minutes, average 2005-2008.

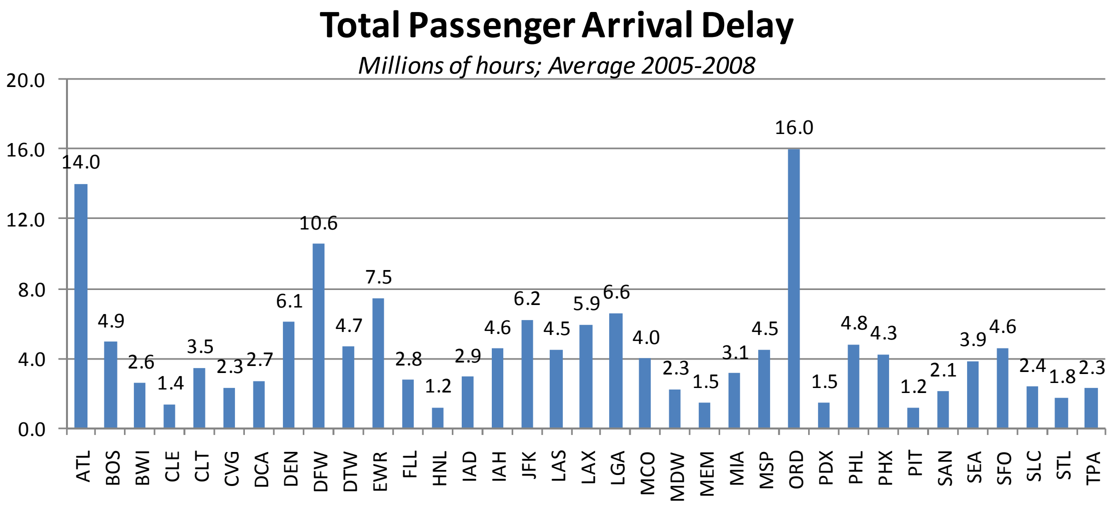

Figure 4.48 - Total annual passenger arrival delay (millions of hours). Average 2005-2008.

4.4.3.1.3 Benchmark Results and Discussion

The results of the operational efficiency benchmarks are presented in Figure 4.49. The results indicate that six airports are fully efficient: ATL, HNL, JFK, LAS, LAX, and SAN. The results indicate that the airports with the lowest performance are TPA, STL, BOS, MCO, and DTW. The full details of the results are provided in Appendix E. This section discusses the implications of these results and investigates the impact of several different factors on the results.

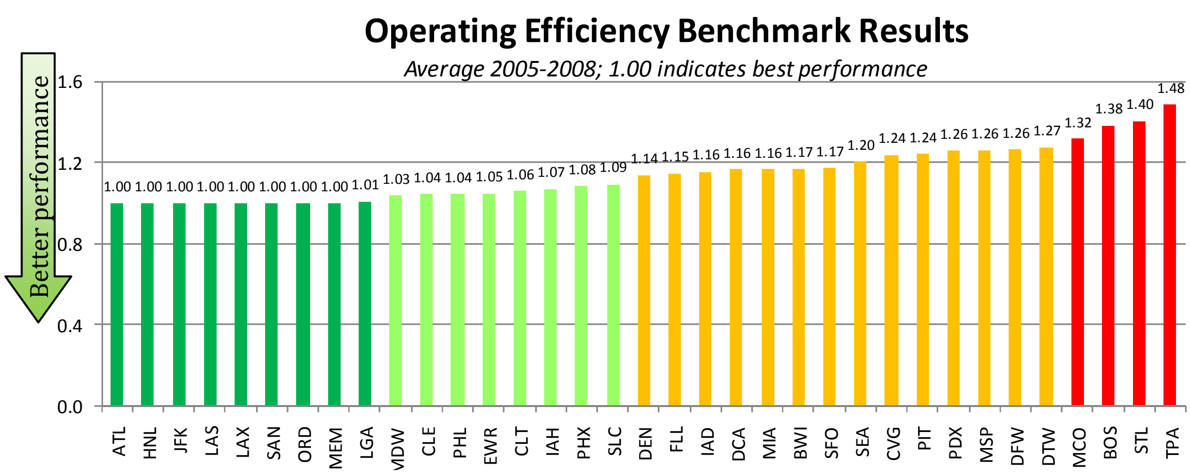

Figure 4.49 - Results from operating efficiency benchmark. Average 2005-2008. 1.00 indicates best performance. Bar coloring is based on a k-means cluster analysis of benchmark results.

4.4.3.1.3.1 Impact of Airport Size on Results

To test the impact of airport size on the results as was done for the original study, the efficiency ranks were computed for small and large airports and a Kruskal-Wallis test was conducted to test the difference between the ranks. A scatter plot of the relationship is presented in Figure 4.50.

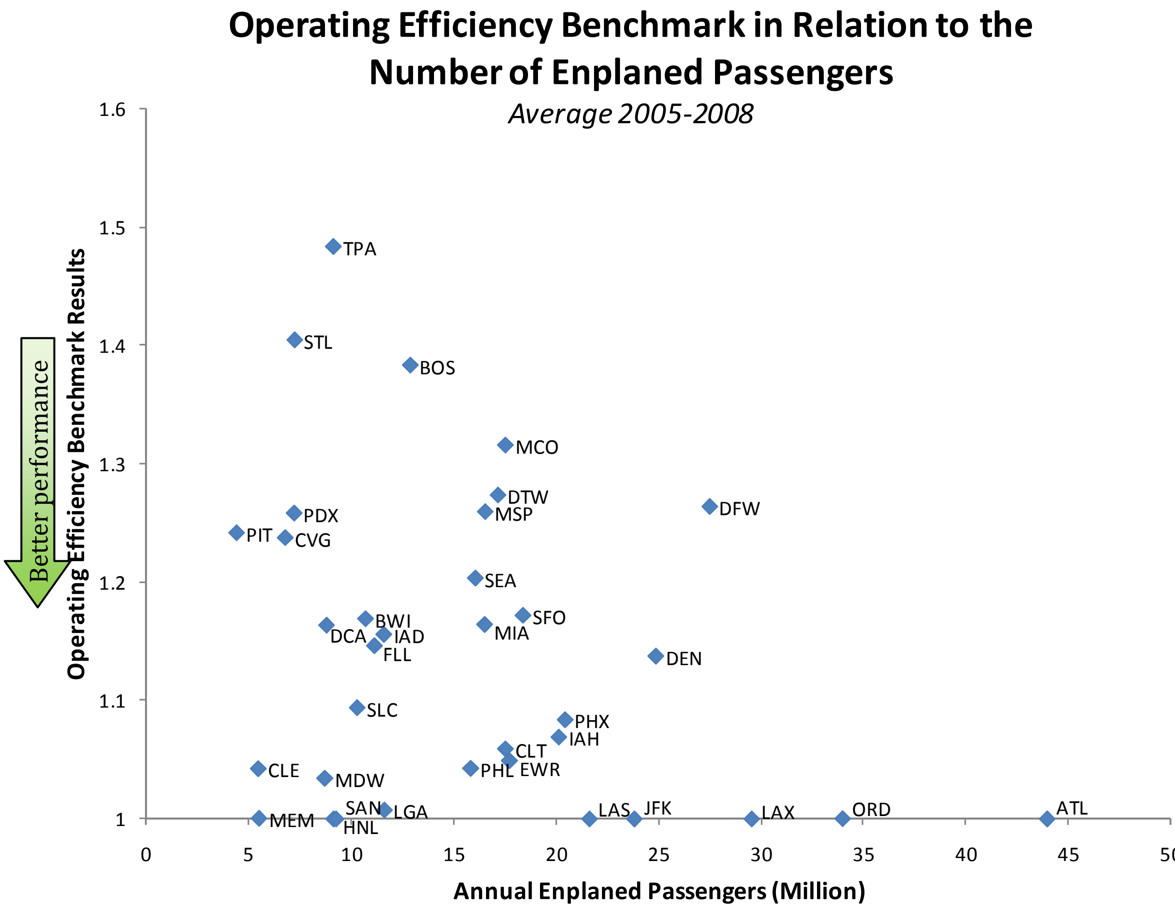

Figure 4.50 – The operating efficiency benchmark results in relation to the number of enplaned passengers. A low efficiency score indicates strong performance

As described in section 4.4.1.3, the airports were grouped by size according to the number of enplaned passengers, with the 17 smaller airports making up the "small" group and the remaining 18 airports categorized as "large". As it is not possible to assume that the efficiency scores are normally distributed, the t-test cannot be used for comparing the performance of the two groups. Instead the Kruskal-Wallis test was performed on the ranks of the efficiency score since the Kruskal-Wallis test does not require that the population have a normal distribution. The mean ranks, χ^2, and significance statistics from the test are presented in Table 4.34.

Table 4.34 – Results from Kruskal-Wallis test of operating efficiency benchmark results ranking based on airport size

Year	Mean rank		Chi-square	Asymptotic significance
	Small	Large		
2005	18.88	14.11	1.798	0.17990
2006	17.53	14.78	0.709	0.39970
2007	18.76	15.28	0.693	0.40530
2008	15.71	16.50	0.041	0.83980

The results indicate no statistically significant difference in the rankings between the two groups of airports for any year in the study. This is in contrast to the test in the original study, where a statistically significant difference in ranks between the two groups was detected for every year.

4.4.3.1.3.2 Impact of Weather Conditions on Results

Past benchmarks have reviewed the impact of weather conditions on airport performance; for instance, (Sarkis 2000) found worse performance for airports operating in the "snow belt". In this study of operating efficiency, the impact of weather conditions is also of interest, and the hypothesis is formulated that airports in poor weather conditions exhibit worse operating efficiency due to reduced airport arrival rates and higher delays.

To test this, data on weather conditions at airports was retrieved from the ASPM database (Federal Aviation Administration 2010d). The database contains observations on the weather conditions at the airports, classifying the weather impact as "none", "minor", "moderate", or "severe" (FAA Office of Aviation Policy and Plans 2005) based on reporting from the Integrated Terminal Weather System (Lincoln Laboratory, Massachusetts Institute of Technology 2010).

Data on the percentage of time between the hours of 07:00 and 22:59 (local time) when each type of weather condition prevailed was collected for the period 2005-2008. The portion of time when the weather conditions were either "moderate" or "severe" is presented for each airport in Figure 4.51.

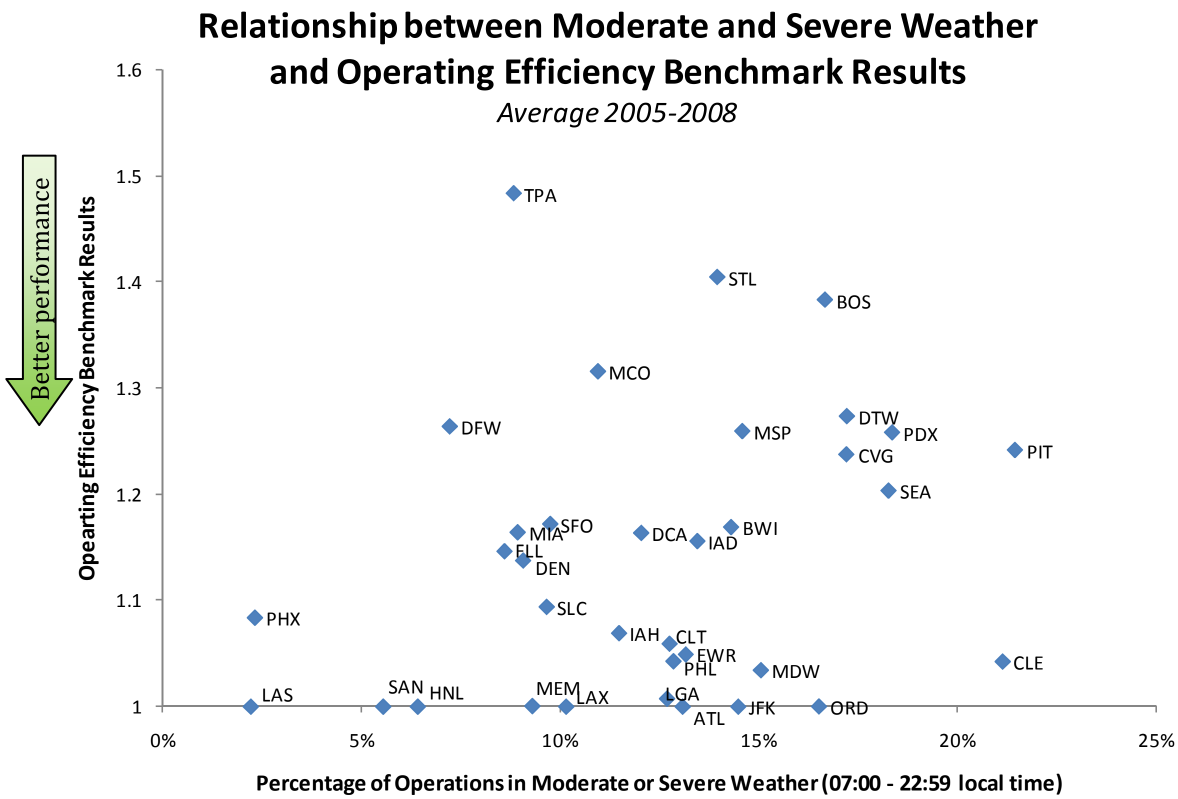

Figure 4.51 - Relationship between moderate and severe weather conditions and the operating efficiency benchmark results; average 2005-2008

To test the hypothesis about the weather impact on results, the Kruskal-Wallis test was applied. Airports were divided into two groups; the 17 airports with the lowest portion of time with moderate or severe weather were classified as "good weather" and the remaining 18 airports were labeled "bad weather". The results of the Kruskal-Wallis test are shown in Table 4.35. The test results do not support the hypothesis that airports in bad weather conditions have worse benchmark performance. This suggests that other factors have a greater impact in determining the operating efficiency of airports.

Table 4.35 - Results from Kruskal-Wallis test of operating efficiency benchmark results ranking based on weather conditions

Years	Mean rank: Good weather	Mean rank: Bad weather	Chi-square	Asymptotic significance
2005-2008	15.88	19.17	1.232	0.26710

4.4.3.2 Investment Quality Benchmark

The investment quality benchmark is computed for only one year. Actual airport credit ratings were available for September 2009, and the objective was to compare the investment quality benchmark to the corresponding time period's credit ratings. Since the analysis was based on one full year's worth of data, 2008 was selected as that would be the most recent year for which a full year's data was available in September 2009. When rates of change were sought, they were based on the change in full-year performance from 2007 to 2008.

4.4.3.2.1 Data Sources

Several data sources were used to assemble each of the parameters for the investment quality component benchmarks:

- **Population growth**: Data on the population of each MSA was gathered from the U.S. Census Bureau (U.S. Census Bureau 2010b). The annual MSA population is estimated by the Census Bureau based on the Census 2000 combined with a number of more recent data sources. The Census Bureau points out that because there is a lag in some of the data sources that complement the Census 2000 data, estimates for older vintages tend to be more accurate than those for more recent vintages (U.S. Census Bureau 2008).
- **Regional GDP growth**: Data on GDP by MSA was obtained from the U.S. government's Bureau of Economic Analysis (BEA) (Bureau of Economic Analysis, U.S. Department of Commerce 2010). The BEA produces annual estimates of the GDP of each of the 366 U.S. MSAs by computing the sum of the GDP originating in all industries in each MSA.
- **O&D passenger percentage**: Data on the level of O&D passenger traffic at each of the airports was computed from the DB1B database

(Bureau of Transportation Statistics 2010c) by summing all passenger itineraries that started or ended at an airport and dividing it by the sum of all enplanements and deplanements at that airport (including both connecting and O&D passengers). This data was only available for domestic traffic.

- **Portion of main carrier's passengers enplaned at this airport**: Data for this factor was computed from the T100 database and included both domestic and international traffic.
- **Portion of OEP-35 passengers**: Data for this factor was computed from the T100 database and included both domestic and international traffic.
- **Debt service coverage ratio**: This data was computed from information retrieved from the FAA's Compliance Activity Tracking System (Federal Aviation Administration 2010a).
- **Non-aeronautical revenue percentage**: This data was computed from information retrieved from the FAA's Compliance Activity Tracking System (Federal Aviation Administration 2010a).

4.4.3.2.2 Summary of Benchmark Parameters

This section presents each of the eight parameters used in the benchmark for 2008. The full details of the parameters are provided in Appendix E.

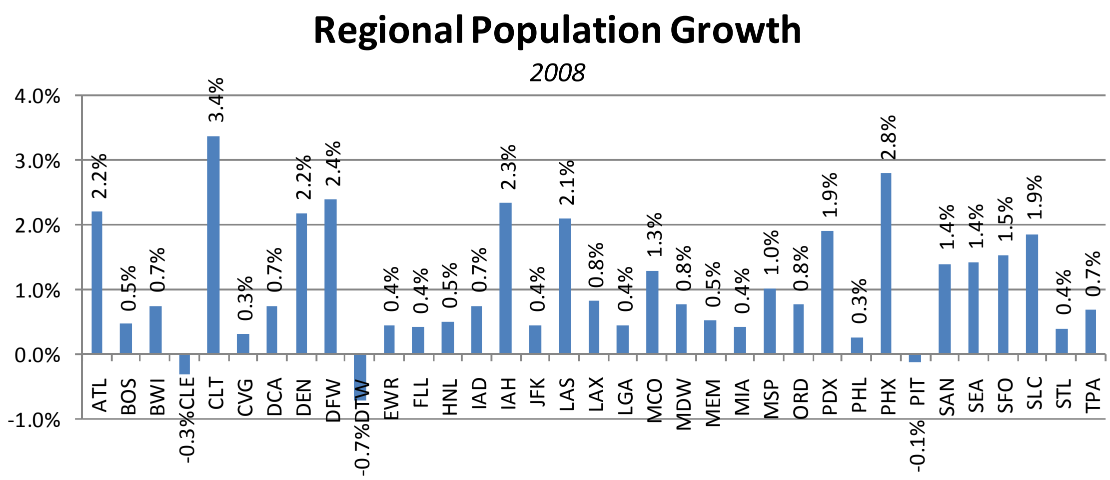

Figure 4.52 - Regional population growth, 2008. Computed as the change from 2007 to 2008.

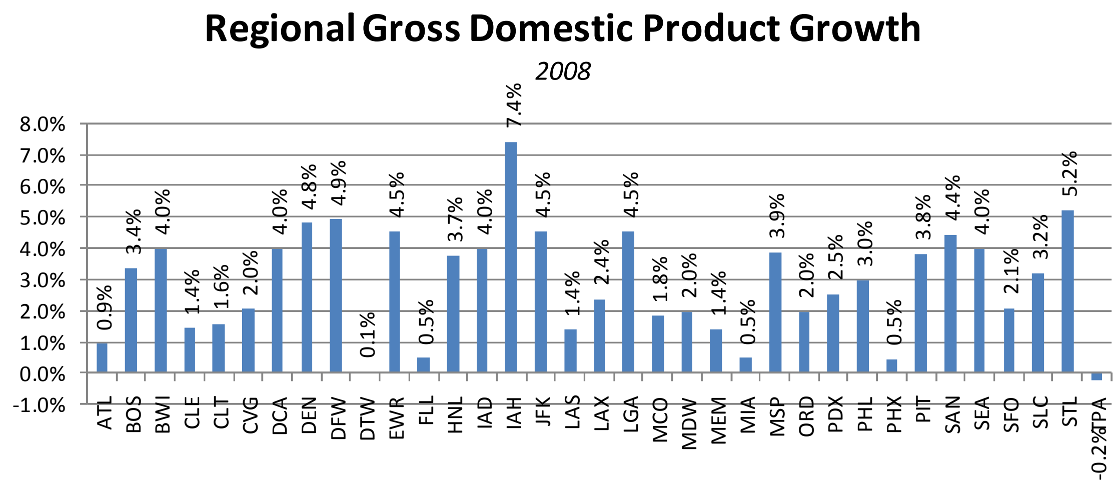

Figure 4.53 - Regional Gross Domestic Product growth. Computed as the change from 2007 to 2008.

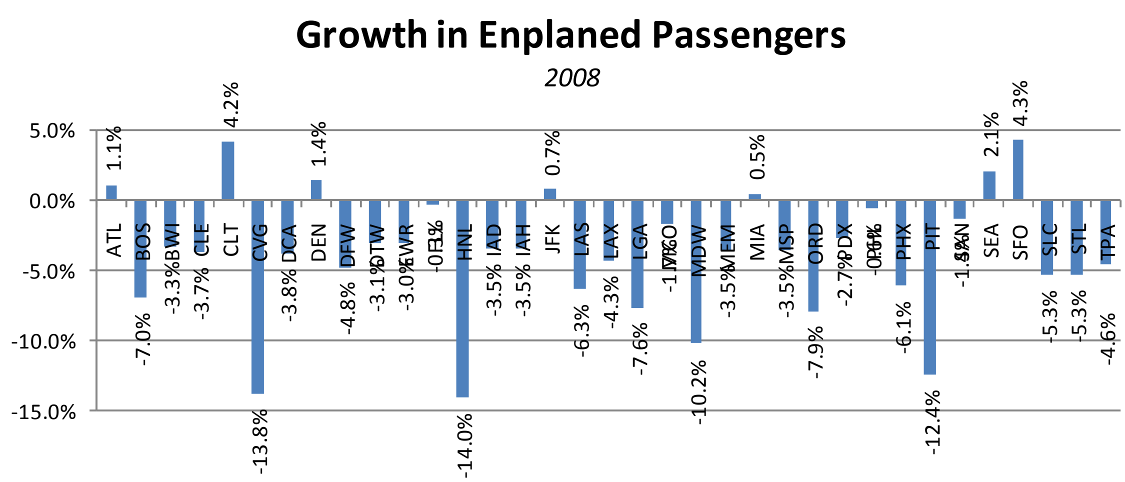

Figure 4.54 - Growth in enplaned passengers, 2008. Computed as the change from 2007 to 2008.

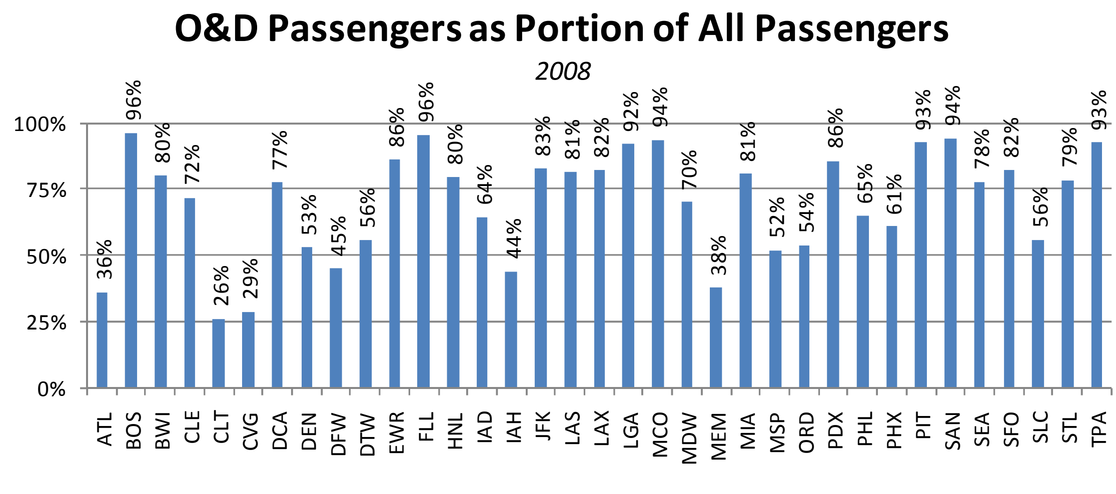

Figure 4.55 - O&D passengers as portion of all passengers.

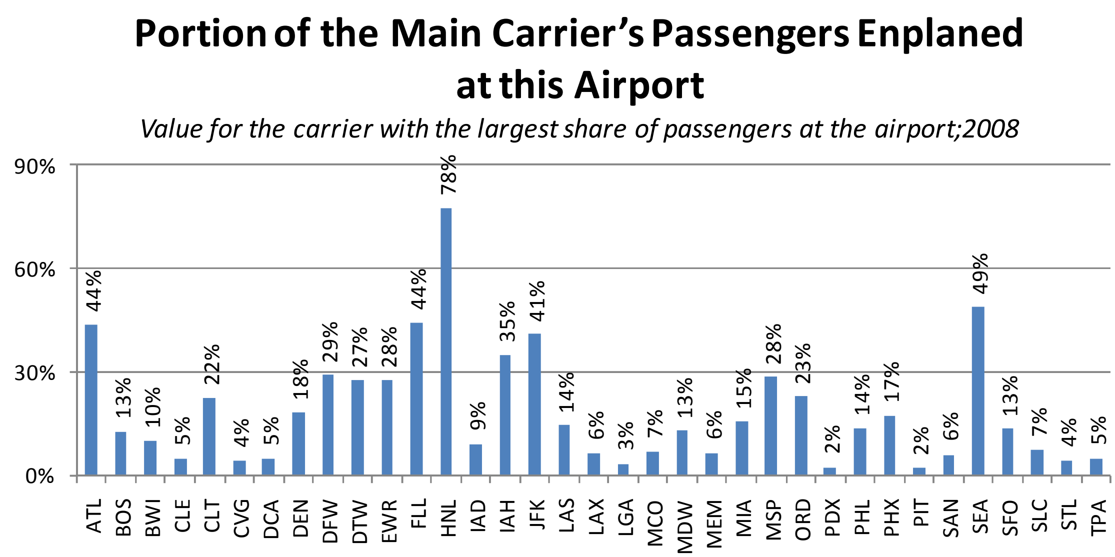

Figure 4.56 - Portion of the passengers that were enplaned at this airport for the carrier with the the largest share of this airport's passengers.

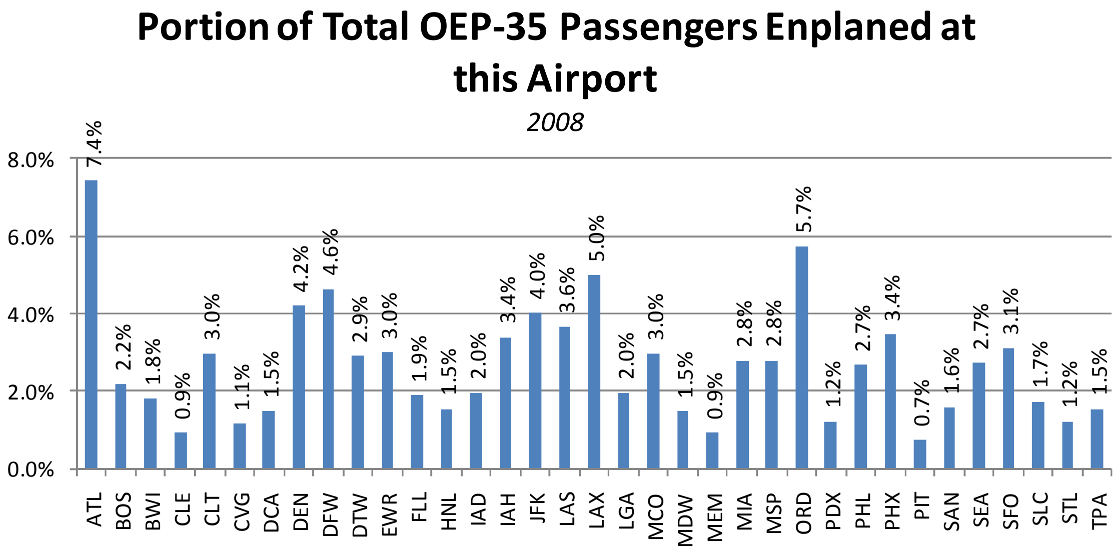

Figure 4.57 - Portion of total OEP-35 passengers (including both international and domestic) that were enplaned at this airport.

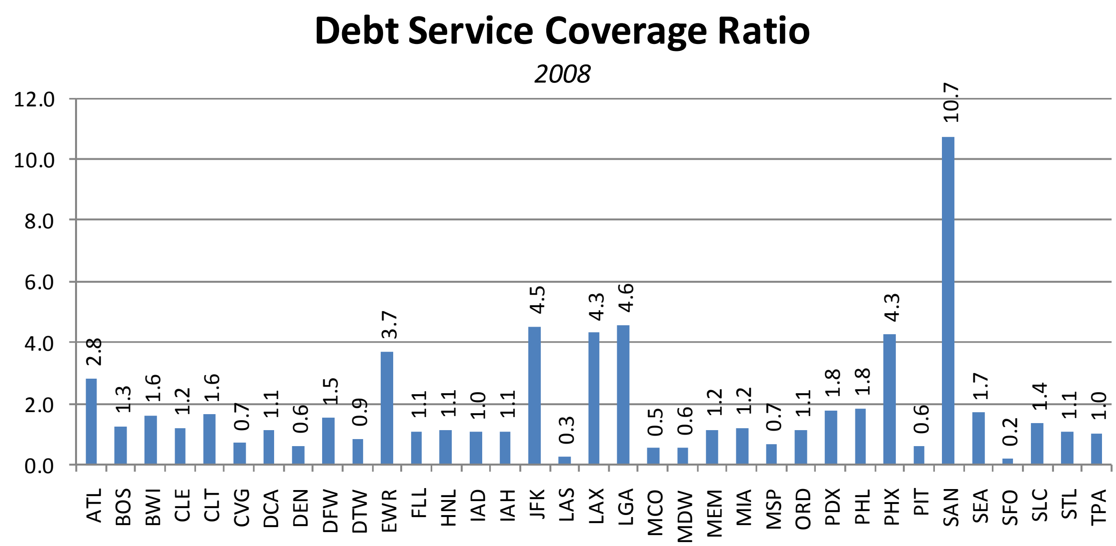

Figure 4.58 - Debt service coverage ratio. The formula used for computing the ratio is provided in Figure 4.44. A high debt service coverage ratio is desirable.

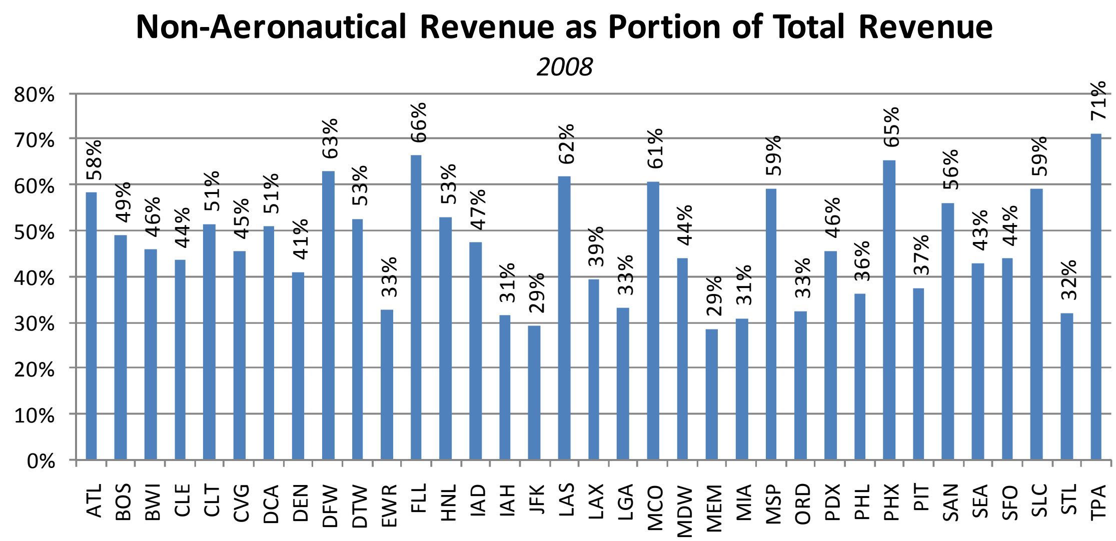

Figure 4.59 - Non-aeronautical revenue as percentage of total revenue.

4.4.3.2.3 Benchmark Results and Discussion

The benchmark results for the three components of the investment quality benchmark are presented in Figure 4.60, Figure 4.61, and Figure 4.62. The DEA model used in this study generates efficiency scores where a low value indicates good performance. The full details of the results are provided in Appendix E. The results of each benchmark show little overlap between the airports on the efficient frontier: For the regional growth benchmark, CLT and IAH are fully efficient and far outpace all other airports. For the air service benchmark, a large number of airports are on the efficient frontier: ATL, BOS, CLT, FLL, HNL, JFK, LAX, MCO, SEA, and SFO. Finally, in the financial factors benchmark, SAN and TPA make up the efficient frontier, and along with the nearly-efficient PHX, they have far better performance than all other airports. The following subsections discuss the impact of several factors on these results.

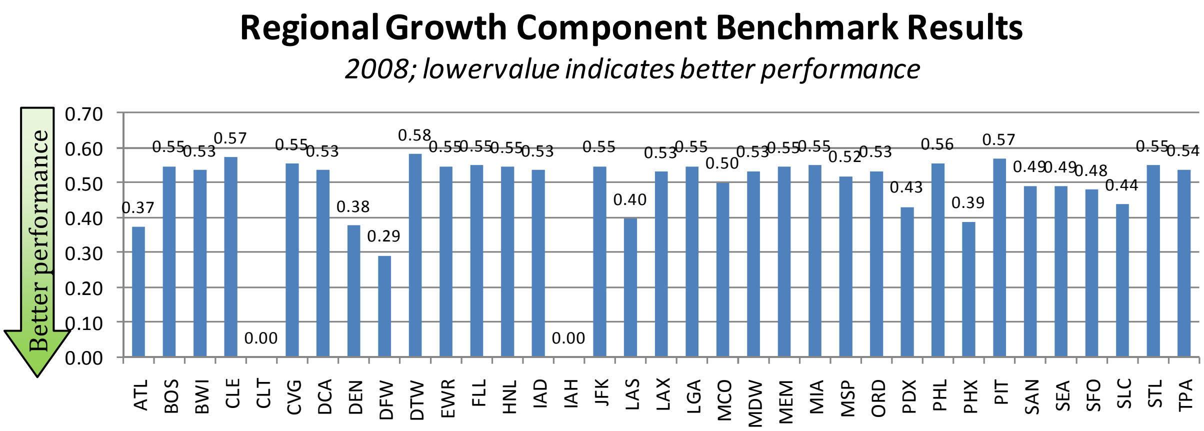

Figure 4.60 - Regional growth component benchmark results for 2008. A lower value indicates better performance.

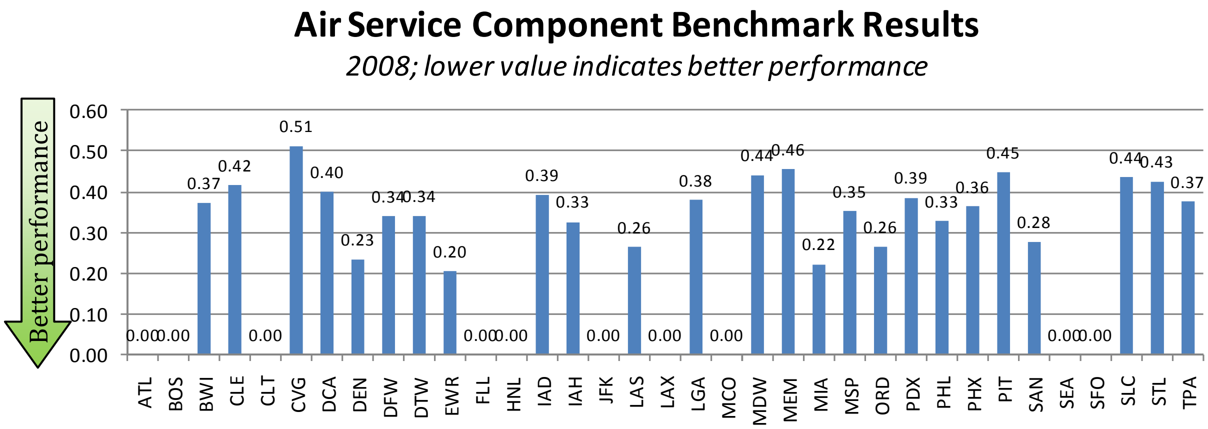

Figure 4.61 – Air service component benchmark results for 2008. A lower value indicates better performance.

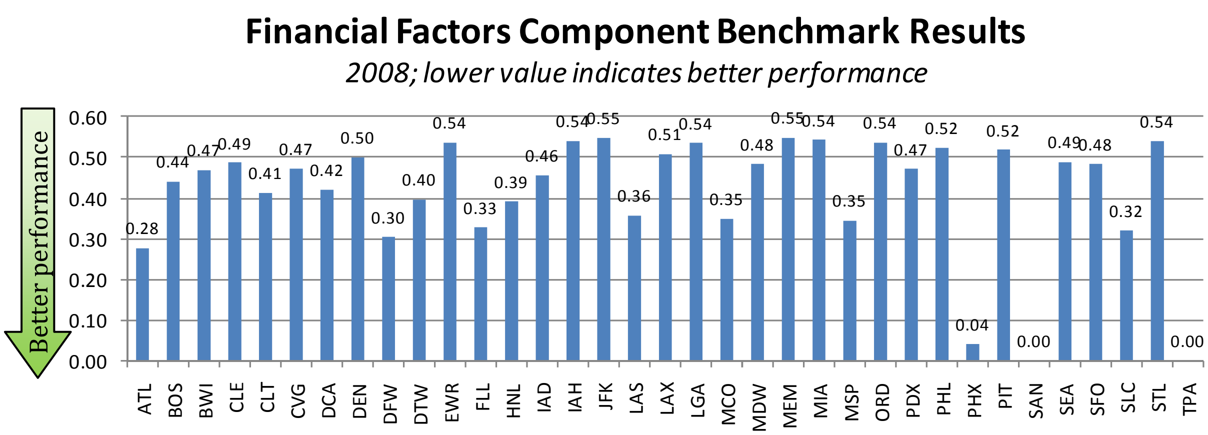

Figure 4.62 – Financial factors component benchmark results for 2008. A lower value indicates better performance.

4.4.3.2.3.1 Impact of Airport Size on Results

As with the original benchmark, of interest is the impact of the airport size on benchmark results. The regional growth benchmark is not reviewed for any impact of airport size since no cause-and-effect relationship hypothesis can be described. However, the level of air service and financial factors benchmarks are both examined for any impact of airport size in this section. The two benchmark results are plotted against the volume of enplaned passengers in Figure 4.63 and Figure 4.64.

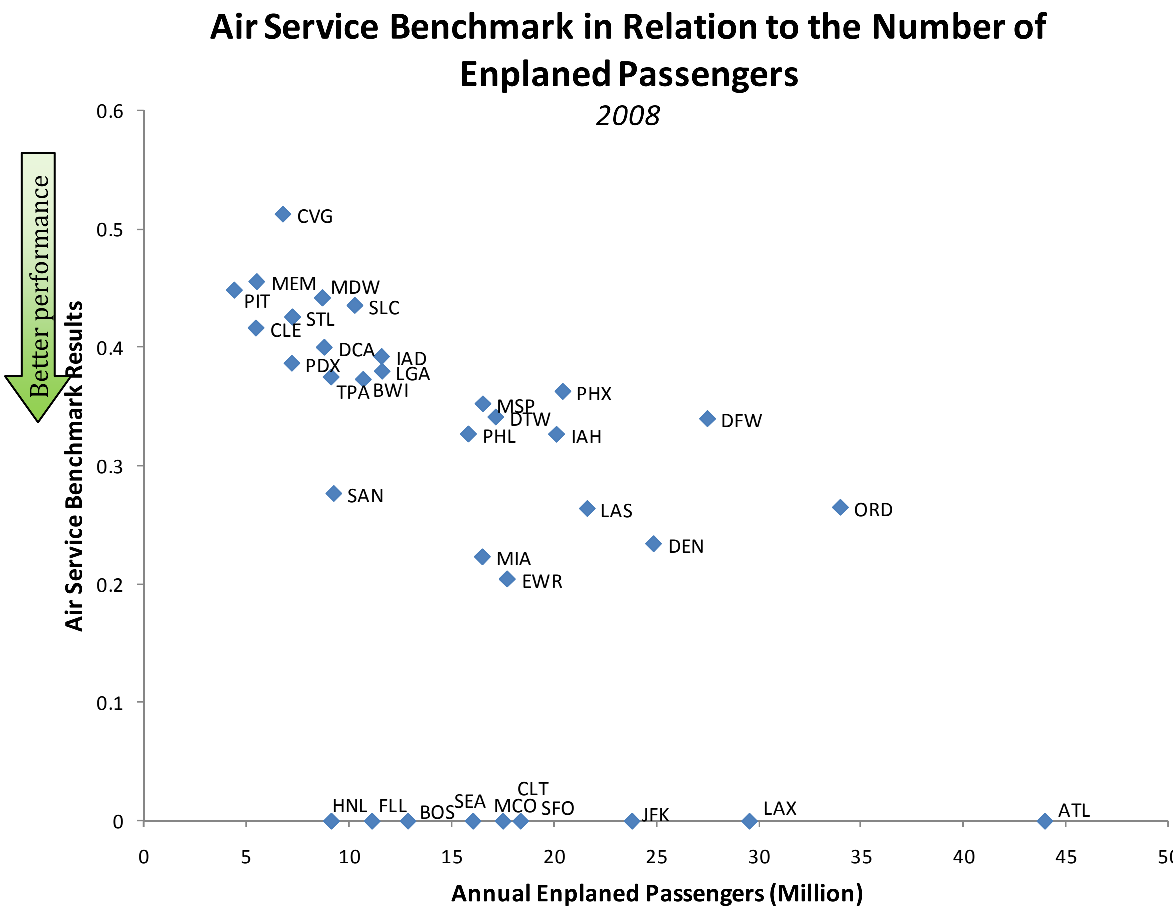

Figure 4.63 - The relationship between the volume of passengers and the results of the air service benchmark. 2008.

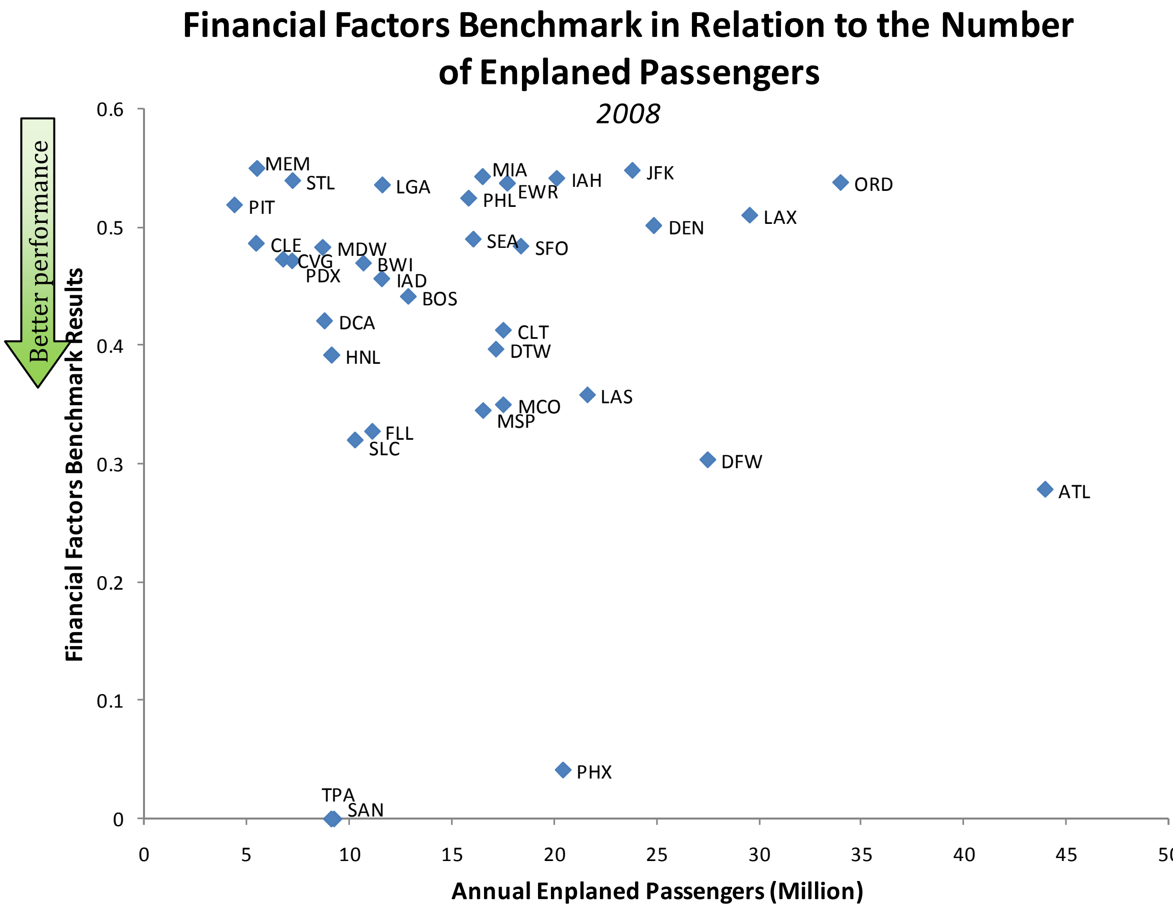

Figure 4.64 - The relationship between the volume of passengers and the results of the financial factors benchmark. 2008.

The relationships were tested using the same Kruskal-Wallis test as in section 4.4.1.3 and 4.4.3.1.3.1 since the benchmark results could not be assumed to be normally distributed. The results of the Kruskal-Wallis test are shown in Table 4.36 and Table 4.37.

Table 4.36 - Results from Kruskal-Wallis test of air service benchmark results ranking for small and large airports

Year	Mean rank		Chi-square	Asymptotic significance
	Small	Large		
2008	23.29	10.50	11.945	0.00055

Table 4.37 - Results from Kruskal-Wallis test of financial factors benchmark results ranking for small and large airports

Year	Mean rank		Chi-square	Asymptotic significance
	Small	Large		
2008	16.88	19.00	0.353	0.55240

The results of the Kruskal-Wallis test show a statistically significant difference at the 95% level for the air service benchmark, with large airports achieving better performance than small airports. It should be noted that two of the factors used in this benchmark (the portion of the dominant carrier's passengers enplaned at this airport, and the portion of the OEP-35 passengers enplaned at this airport) are not independent of the volume of passengers. This is an issue that cannot be eliminated since any measure of the size of an airport's throughput will be positively correlated with the volume of passengers carried.

For the financial factors benchmark, no statistically significant difference between the two groups of airports was detected, indicating that no statement can be made about the financial factors performance of airports in relation to their size.

4.4.3.2.3.2 Impact of the Level of Connecting Traffic

The level of connecting traffic is of interest as it relates to the level of financial factors performance. It is not tested in for its impact on regional growth rates or on the level of air service. Testing the impact of the degree of connecting traffic on financial factors performance gives an indication of whether or not it is attractive for an airport to serve a high degree of connecting passengers.

As in section 4.4.3.2.1, data on the level of domestic connecting traffic at each of the airports was computed from the DB1B database (Bureau of Transportation Statistics 2010c) by summing all passenger itineraries that started or ended at an airport and dividing it by the sum of all enplanements and deplanements at that airport (including both connecting and O&D passengers). Figure 4.65 shows the relationship between the level of connecting traffic and the financial factors benchmark results.

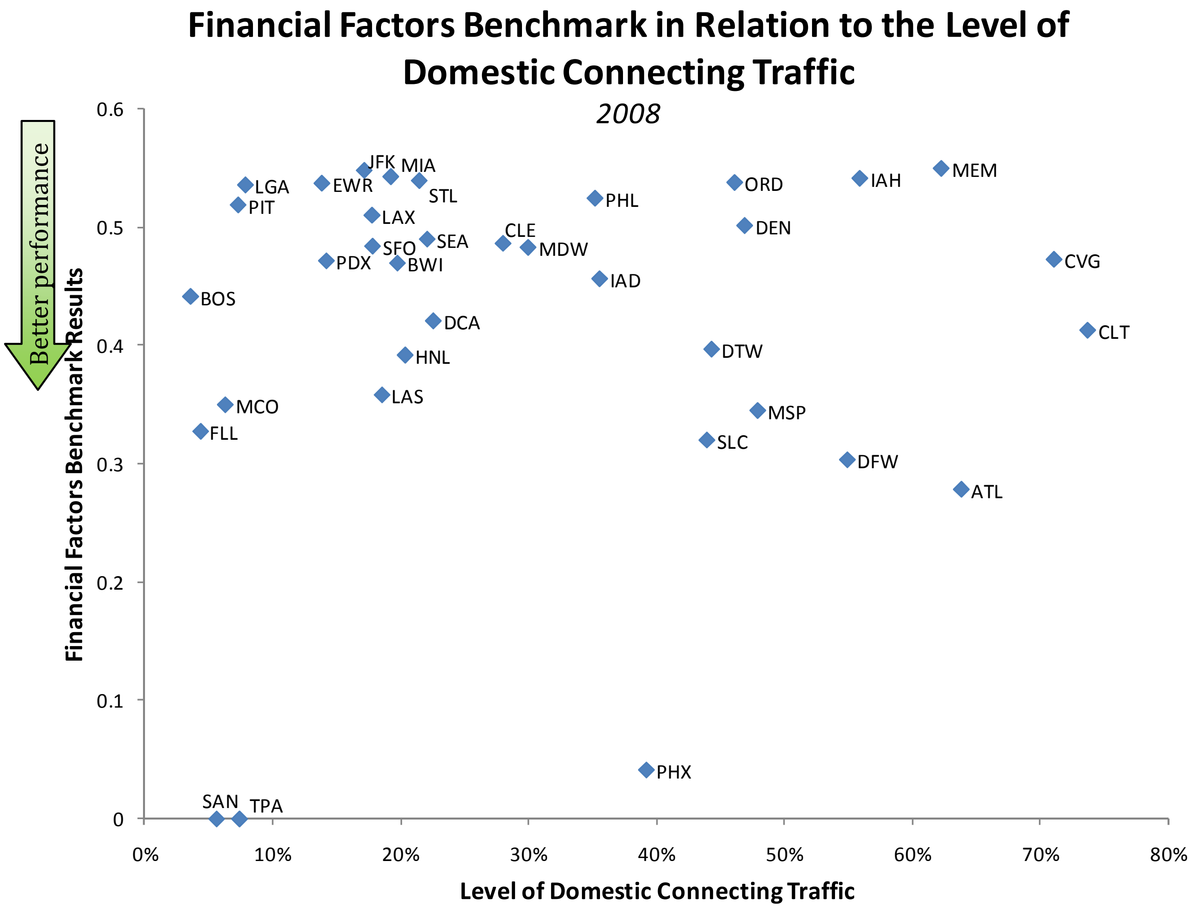

Figure 4.65 - The relationship between the level of domestic connecting traffic and financial factors benchmark results, 2008. A low benchmark result score indicates strong performance

The Kruskal-Wallis test on this relationship was conducted by dividing the airports into two nearly even-sized groups: 17 airports with low levels of connecting traffic (≤ 22% connecting traffic) and 18 airports with high levels of connecting traffic (> 22% connecting traffic). The results of the Kruskal-Wallis test are shown in Table 4.38.

Table 4.38 - Results from Kruskal-Wallis test of financial factors benchmark results ranking for low and high degrees of connecting traffic.

Year	Mean rank		Chi-square	Asymptotic significance
	Low	High		
2008	18.59	17.39	0.132	0.71650

It should be noted that a limitation of these results is the fact that the level of connecting traffic only includes domestic traffic due to limitations in the availability of data. The results would potentially be different if connecting international traffic were included.

4.4.3.2.4 Relationship of Benchmark Results to Airport Credit Ratings

Airport credit ratings are issued by the credit rating agencies Moody's, Fitch, and Standard & Poor's, and data was obtained for the Fitch Ratings evaluation of airport debt for September 2009 (Lehman et al. 2009). Ratings are on a scale from AAA (highest) to D (lowest), with AAA through BBB being considered "investment grade" and BB through D are considered "speculative grade" (Fitch Ratings 2010).

The objective of this section is to study the relationship between the investment quality benchmark and these actual credit ratings. The credit ratings for the airports being studied are shown in Table 4.39. Data was not available for all airports in the benchmark since some airports are not rated by Fitch Ratings. Airport debt is sometimes structured in different slices ("tranches") of debt, with more senior slices having precedence in being repaid before other slices. The most senior slice will be assigned the highest credit rating. This structure varies by airport, which causes some comparability issues when comparing credit ratings; the credit rating of each airport is determined not only by the airport's performance but also by the way that the airport's debt is structured. What is presented in Table 4.39 is the rating of the most senior slice of debt for each airport.

Table 4.39 - Credit rating of most senior airport liens (Lehman et al. 2009)

Airport	Rating of most senior lien
ATL	A+
BOS	AA
BWI	A
CLE	A
CLT	A+
CVG	A-
DCA	AA
DEN	A+
DFW	AA-
DTW	A
EWR	AA-

Airport	Rating of most senior lien
FLL	A+
HNL	A
IAD	AA
IAH	A+
JFK	AA-
LAS	A+
LAX	AA
LGA	AA-
MCO	AA-
MDW	A+
MEM	A+
MIA	A
MSP	AA-
ORD	AA+
PDX	N/A
PHL	A
PHX	N/A
PIT	BBB+
SAN	A+
SEA	AA
SFO	A
SLC	N/A
STL	BBB
TPA	AA-

To compare the benchmark results with the actual airport credit ratings, a Spearman correlation test was used, with Spearman correlation being the appropriate method since two different ranks are being compared. To make the test possible, several design choices were made:

1. The credit ratings were translated from their letter grades to a ranking. The assumption was made that the letter grades progress on a linear scale, such that the gaps between AAA, AAA-, AA+, AA, AA-, etc., are proportionally sized. In the analysis, AAA was assigned the value 1, AAA- was assigned 2, etc.
2. Since the results from the benchmark were computed as three separate components, they had to be combined into one overall assessment of performance. However, the three scores were computed on different scales (i.e. a score of 0.5 in the air service benchmark does not have the same meaning as 0.5 for the regional growth benchmark) so they could not be directly combined. Instead, the benchmark results were converted to a ranking, with the best airports being assigned a rank of 1 in the benchmark and all others higher rank values than that.
3. The mean value of the three component benchmark rankings was used for the Spearman correlation test. The data for this test is plotted in Figure 4.66.

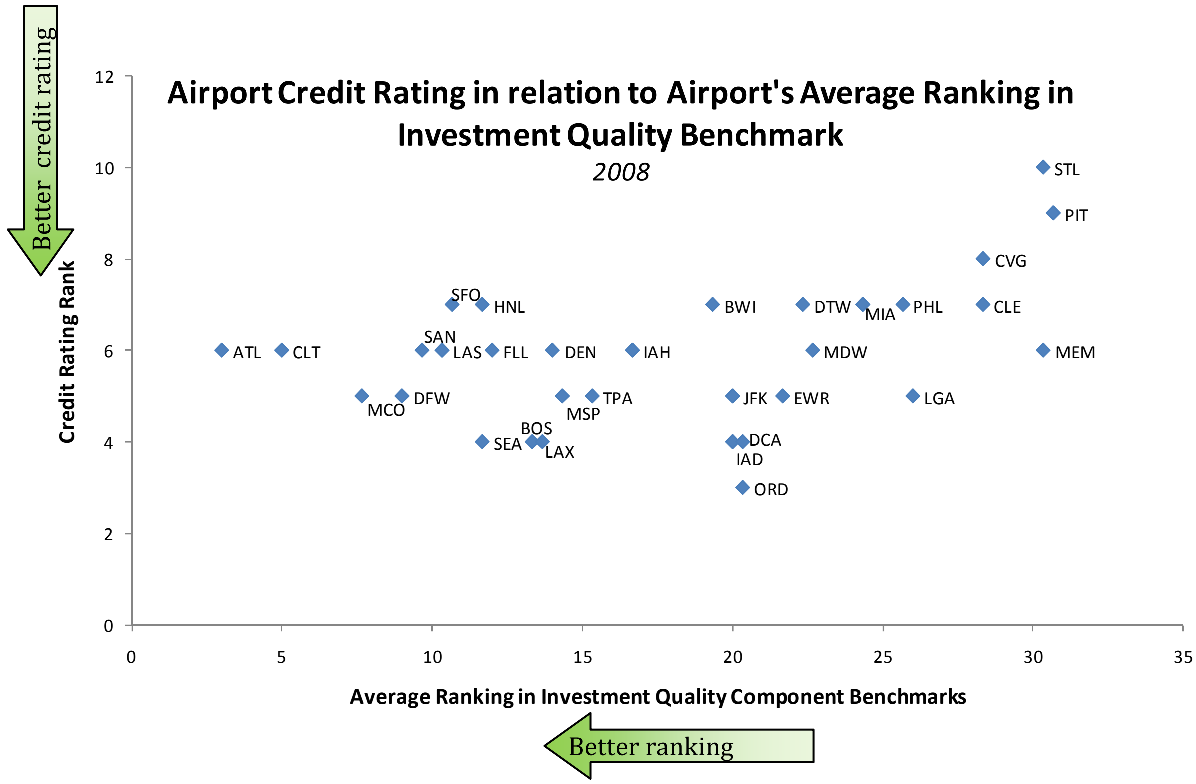

Figure 4.66 – Airport credit rating as a function of the airport's average ranking in the investment quality benchmark, 2008.

In the test, the Spearman rank correlation coefficient ρ = 0.37, with p = 0.039, indicating that the correlation is significant at the 95% level.

The somewhat low correlation coefficient can be attributed to factors such as:

1. The credit rating of the most senior slice of debt varies depending on not only the airport's performance but also on the structure of its debt, as described earlier in this section.
2. The factors included in the investment quality benchmark are only a subset of the factors considered by credit ratings agencies.

4.4.4 Comparison of Original and New Benchmark Results

To gauge the relationship of the original benchmark results with the two new component benchmarks, each airport's ranking in the three benchmarks (the investment quality component benchmark rankings were averages in the same way as in section 4.4.3.2.4) were used. Since the investment quality benchmark was only computed for 2008, only 2008 values were used in the comparison. This comparison is presented in Figure 4.67.

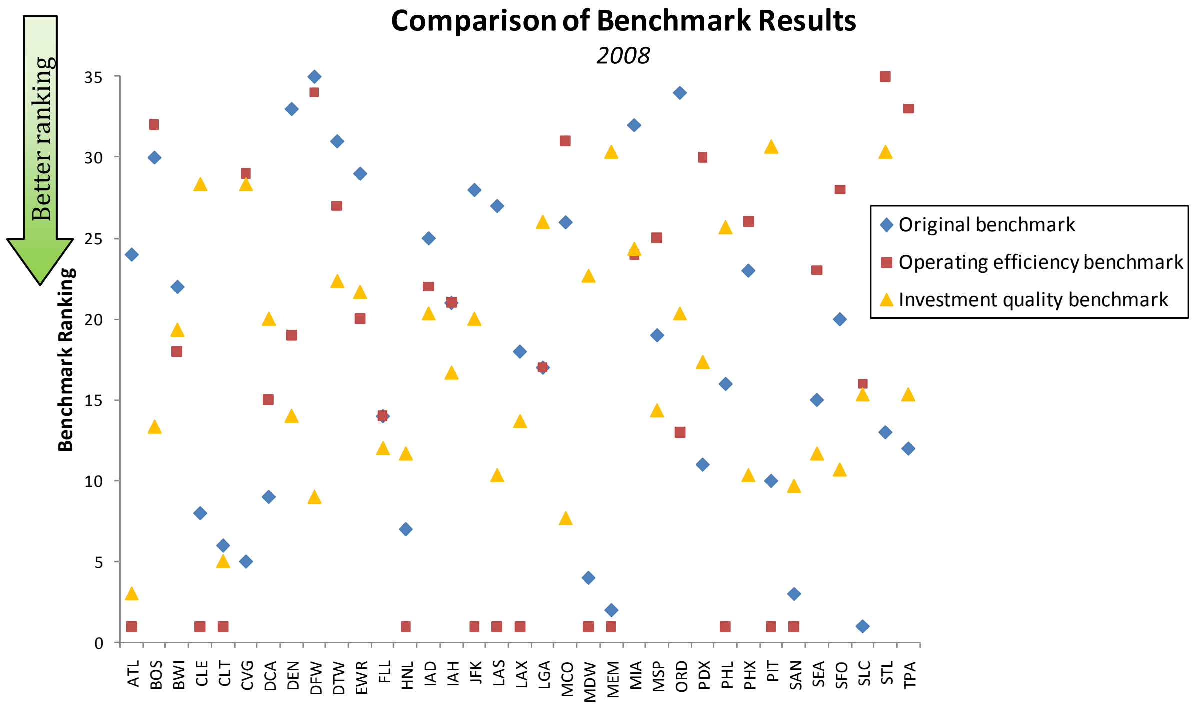

Figure 4.67 - Comparison of ranking of results from original benchmark, operating efficiency benchmark, and investment quality benchmark, 2008.

As Figure 4.67 suggests, the two new benchmark results have a low degree of correlation with the original benchmark. The Spearman rank correlation coefficients are presented in

Table 4.40.

Table 4.40 - Spearman rank correlation coefficients for test of correlation between benchmark result rankings

	Original benchmark	Operating efficiency benchmark	Investment quality benchmark
Original benchmark			
Operating efficiency benchmark	$\rho =$0.34 p = 0.045		
Investment quality benchmark	$\rho =$-0.19 p = 0.268	$\rho =$-0.06 p = 0.751	

The correlation coefficient magnitudes are all low, and only the correlation between the original benchmark and the operating efficiency benchmark is significant at the 95% level. This low level of correlation between the results of the original benchmark and the two new component benchmarks points most importantly to the need for the type of systematic benchmarking methodology presented in this dissertation; if no motivation is provided for the selection of benchmark metrics and for the choice of benchmark model, readers should be skeptical of the results of a benchmark.

Beyond the general criticism of the original study's lack of structure and motivations in its design, three specific weaknesses were identified with the original study:

1. Some of the performance data used in the original benchmark was not appropriate for use in the benchmark without further adjustments or processing due to poor comparability between the airports. For example, the aeronautical revenues and the operating costs exhibited wide variations between the different airports to a degree that cannot entirely be explained by airports' different sizes. This is exemplified by a comparison of ATL and JFK: JFK reports operating costs and aeronautical revenues that were both about five times higher than those of ATL, even though ATL has a higher passenger volume and number of aircraft movements. This difference likely stems from factors such as differences in services provided by the airport (Neufville & Odoni 2003, p. 274).
2. The use of the percentage of on-time performance in the study was not appropriate since this parameter does not scale with the size of other inputs and outputs. If the factor had instead been computed as the number of flights arriving on time, or the total number of delay

minutes, it could have been included since both of those factors do scale with the size of the operation[11].

3. The analysis used too many parameters in relation to the number of DMUs. A total of 4 inputs and 6 outputs were used for the 35 DMUs, which violates the rule of thumb guidance of (R. G. Dyson et al. 2001) that the number of DMUs should be at least twice as large as the product of the number of inputs and outputs parameters combined, since 35 ≱ 2 * 4 * 6. This would have caused too a high portion of DMUs being deemed fully efficient, but rather than scoping down the number of variables (for instance, aeronautical and non-aeronautical revenues could have been combined and the number of air carrier operations could have been summed with the number of other operations), the authors used an artificial approach of a super-

---

[11] It should be noted that percentage parameters were used in the investment quality benchmark. This was possible since the investment quality benchmark only used percentages and ratios that do not scale with the size of the operation. Where the use of percentages becomes a problem is when they are mixed with other parameters that do scale with size.

efficient DMU to force full ranking of all DMUs even though no motivation existed why no ties could exist.

4.4.5 Conclusions

This section presented the results of applying the new benchmark methodology on a benchmark from the literature to compare the consistency of the results from a redesigned benchmark with those of the original benchmark.

The new benchmarks presented a number of findings about specific airports' performance in terms of operational efficiency and their level of investment quality. It also presented evidence about the existence of a link between efficiency and the size of the airport for some components of the new benchmark, and the lack thereof for others.

However, at a methodological level, the redesigned benchmark presents two key conclusions:

1. The benchmark results are heavily dependent on the design of the benchmark, in terms of the stakeholder goals addressed, the selection of performance metrics, and determination of the benchmark model

to use for computing the results. Changes in assumptions can lead to differences in the benchmark results.

2. As a consequence of the importance of the process of designing the benchmark, the need is underlined for a systematic approach to selecting stakeholder goals and performance metrics, and to selecting the benchmark model to be used. Readers of benchmark results should review the benchmark design process before accepting the findings of a benchmark.

5 Chapter 5: Conclusions and Future Work

The conclusions of the dissertation are of two types. The first are conclusions that relate to the alternative method for airport benchmarking. These are conclusions that relate to how future benchmarks of airport performance should be conducted. These conclusions are presented in the first three subsections. The second type is a set of practical conclusions about airport performance benchmarking that apply to policymakers. These conclusions are presented in the fourth subsection. The final subsection describes directions for future research.

5.1 Airports as Utilities

Major U.S. airports serve a utility-like function for the regions in which they are located, as described in section 2.1. The airports exist in a form of economic symbiosis with their surrounding regions, whereby growth in the local economy fuels increased demand for air travel, and the activities at the airport in turn fuels further regional economic growth. Operating in a monopolistic or oligopolistic environment, airports are barred from generating a profit for the cities, counties, and other entities that own them. Instead of generating a financial surplus, airport performance goals are defined by the stakeholders' objectives for the airports.

This dissertation presented a model of these stakeholders, their interrelationships, and their objectives for the airport, highlighting some objectives which are in alignment with each other and others which are in conflict. The stakeholder model also indicated that many of the factors on which airports are evaluated by stakeholders are not within the direct control of airport management, but are rather managed by the service providers with which the airport organization collaborates to deliver a complete airport service.

The airport stakeholders include the airport organization; the local government, residents, and businesses; air carriers; service providers and concessionaires; investors; and the federal government. These stakeholders' goals for the airport vary widely, and include ensuring high levels of air service; generating growth in passenger volumes; minimizing costs and maximizing revenues, with a focus on non-aeronautical revenues; and minimizing noise and emissions.

Since traditional profit metrics do no gauge how well airports are meeting these stakeholder goals, alternate evaluation methods are needed. Comparative benchmarking is a method which allows for evaluation of performance in multiple dimensions, and it is an important tool for both the airport organization and the airport's stakeholders. Benchmarking is presently in use for evaluating airports, but current benchmarking methods have several shortcomings, which are summarized in the following subsection.

5.2 Limitations of Benchmarking Methods

In analyzing the methodologies used in past benchmarks, the dissertation has demonstrated that several areas of methodological misalignment exist. These misalignments include:

- **Ambivalence about the stakeholder goals being reflected**: Past benchmarks have lacked anchoring of benchmark parameter selection in a model of stakeholders and their goals. This has caused a lack of clarity about what the benchmark results mean since there is little or no motivation for the selection of the particular metrics being used.
- **Lack of motivation of DEA model selection**: Past benchmarks have not systematically addressed why a particular DEA model was selected over another to compute the benchmark results. Research described in section 2.3.2 shows that model selection impacts

benchmark results, sometimes to a drastic level. The analysis in section 4.1.2 shows that the highest degrees of methodological misalignment exist in:

- Aggregation function selection
- Modeling of returns to scale
- Use of integer constraints in DEA modeling
- Treatment of multiple-timespan analysis

These findings call into question the results of past benchmarks on two accounts: The first is the limited motivation for selection of performance metrics. This limitation results in a lack of determination about what the benchmark results mean; are the airports that are identified as the best performing airports in fact excelling in ways that are important to its stakeholders? The second area of concern is the limited motivation for the selection of a particular DEA model. This limitation results in limited validity of the benchmark scores that were computed; if a different model had been selected, the benchmark results would likely not have been the same.

These limitations indicate the need for a comprehensive, systematic approach to airport benchmarking. The following section provides a summary of the methodology developed in the dissertation to address these shortcomings.

5.3 A Comprehensive Benchmarking Methodology

The dissertation has presented a complete benchmarking methodology, and this section summarizes this methodology. The methodology ensures that benchmark results are reflective of stakeholder goals; it ensures that the underlying assumptions of the benchmarking model used is reflective of the characteristics of the environment being modeled; and it ensures that the conclusions are not only a list of airport rankings but that the results are interpreted and turned into findings that explain why some airports exhibit strong performance and others show poor performance.

The benchmarking methodology is summarized in Figure 5.1.

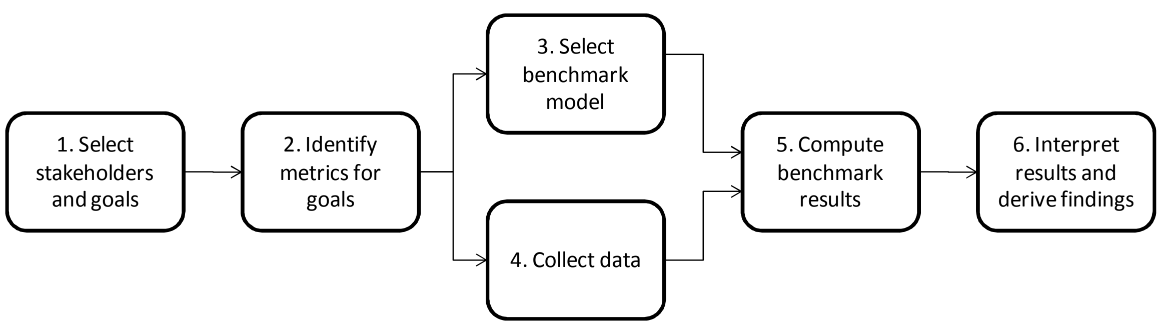

Figure 5.1 - Summary of airport benchmarking methodology

Steps 1, 2, and 3 in the methodology make up the design phase of the benchmark and serve to ensure that the implementation of the benchmark will generate valid results.

Completing steps 1 and 2 in the methodology, in which stakeholder goals are selected and metrics for those goals are defined, creates a benchmark which is reflective of stakeholder interests. These steps are enabled by the model of airport stakeholders and their goals, as described in section 2.1.

Completing step 3 in the methodology, in which the framework and heuristics for selection of a DEA model are applied, creates a benchmark that uses a model whose assumptions are aligned with the domain being modeled. The framework and heuristics are summarized in section 3.3 with full details available in Appendix A.

The benchmark implementation phase consists of steps 4 and 5 in which data is collected and benchmark results are computed, resulting in relative scores of efficiency for each airport.

Finally, step 6 in the methodology represents the third phase of the benchmark, in which the results are analyzed to identify the factors which impact benchmark results. This step ensures that the benchmark results can be used by management, policymakers, and other stakeholders to identify which airports require action to improve performance, and what those actions should be.

In summary, by applying the alternative methodology presented in this dissertation, these limitations described in sections 5.2 can be avoided, and the validity of the benchmark design and results can be assured.

5.4 Implications for Airport Policymakers

Airport policymakers are considered as those responsible in a capacity as lawmaker, as one controlling system resource allocation, or as a regulator. This section describes the implications of the dissertation findings for policymakers.

5.4.1 Summary of Benchmark Results

The graphic in Figure 5.2 presents a visual summary of the benchmark results for policymakers. The coloring in the figure is based on each airport's categorization in the k-means clustering analysis for each benchmark. The airports are listed in descending order based on the aggregated value of their benchmark results, with benchmark results in the first group (dark green) being allocated a value of 1, the second group (light green) a value of 2, the third group (yellow) a value of 3, and the final group (red) a value of 4. In this summary form of the results, ATL ranks at the top of the list, and PIT occupies the bottom. The stack-ranking provided in this list should not be treated as standalone results by themselves, but should rather serve as a starting point of the discussion of the meaning of the results provided in the subsequent sections.

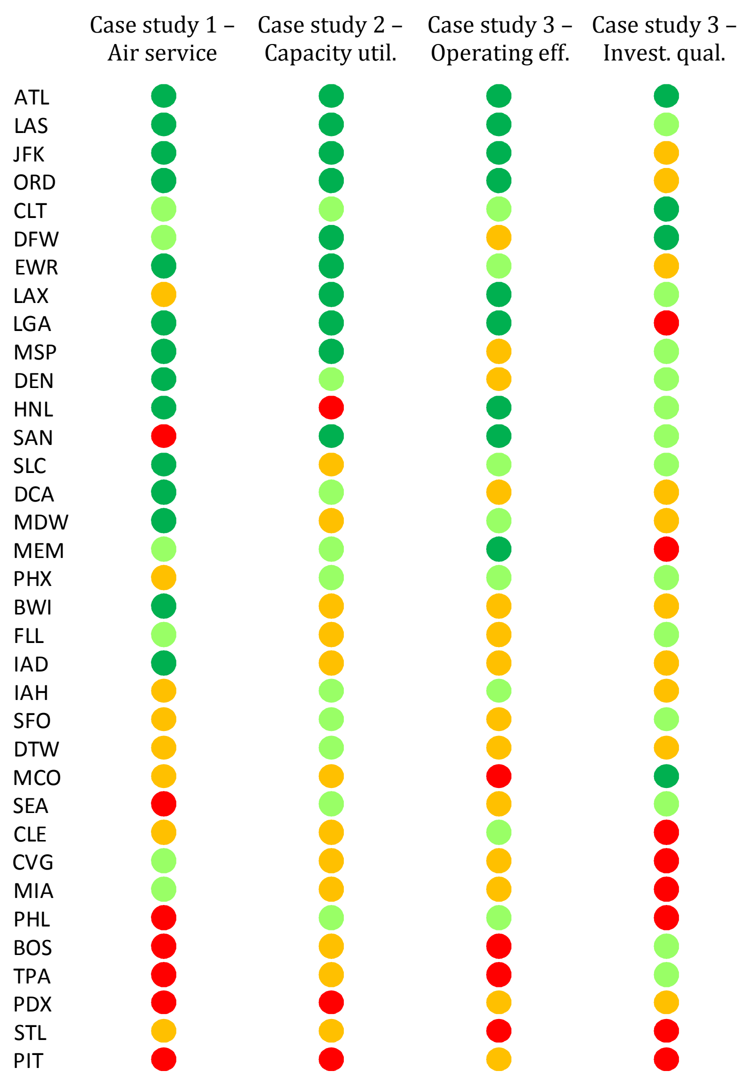

Figure 5.2- Summary of benchmark results. The coloring is based on the k-means clustering of each benchmark's results, ranging from dark green (best performance) to red (worst performance).

5.4.2 Use of Benchmarks by Policymakers

The implications of the analysis of past airport benchmarks in the context of the alternative benchmarking methodology proposed in this dissertation are relevant not only for policymaker stakeholders but for any decision-maker or other consumer of benchmark results: Benchmark results should not be used unless there is clarity about the motivation for selecting a particular set of performance metrics for use in the benchmark. Similarly, the benchmark results should not be used unless there exists a clear motivation for why the benchmark model was selected for computing the benchmark results. If these conditions are not met, benchmark results should not be used for decision-making or any other purpose.

5.4.3 Implications for Decisions about Funding

The benchmark results are conducive for analysis about allocation of federal funding for improvements. As shown in section 2.1.2.1, the federal government's Airport Improvement Program (AIP) provides about 18% of the capital funds for improvements that include enhancements of capacity, safety, and other aspects of airport infrastructure.

As described in section 2.1.3.2.10, the AIP is funded through user fees and fuel taxes. The funds can be used for projects that "support aircraft operations including runways, taxiways, aprons, noise abatement, land purchase, and safety, emergency or snow removal equipment" (Kirk 2003, p. 3). Since Fiscal Year 2001, AIP grants have exceeded $3 billion annually (FAA 2008, p. 69).

The AIP funds are distributed to passenger, cargo, and general aviation airports, and its funds fall in two categories (Kirk 2003, pp. 6-7):

1. **Formula funds**: Formula funds (also known as "apportionments") are apportioned according to formulas based on the volume of throughput (e.g. enplaned passengers) and location. The formulas vary depending on the type of airport.
2. **Discretionary funds**: Discretionary funds are approved by the FAA and are distributed based on factors such as project priority and congressional mandates. Although it is not the sole determinant factor, project selections are based on a project's score in the National Priority Rating (NPR) equation, which assigns projects a rating from 0

to 100 (high)[12], indicating their level of alignment with agency goals (Federal Aviation Administration 2000, p. 5). Projects with safety and security purposes receive higher ratings than those focused on capacity (Dillingham 2000, p. 32).

By studying the allocation of the AIP funds among the airports in the benchmarks presented in the dissertation, the question of whether funds are appropriated to fulfill the greatest needs can be studied.

Figure 5.3, Figure 5.5, and Figure 5.6 show the total AIP spending for fiscal years 2005-2009 on improvements at the airports included in the benchmark, in relation to results from some of the benchmarks. The AIP funding data was gathered from (Federal Aviation Administration 2010b). For the case study of the

---

[12] NPR = .25P*(A+1.4P+C+1.2T), where A is a value representing the type and size of the airport; P represents the project purpose code, e.g. capacity improvement (note that this value is used twice in the formula); C represents the physical component (e.g. runway); and T represents the type of work being done (e.g. extension) (Federal Aviation Administration 2000) (pp. 5-6).

level of air service by region, the AIP spending is summed for all of the airports in a region.

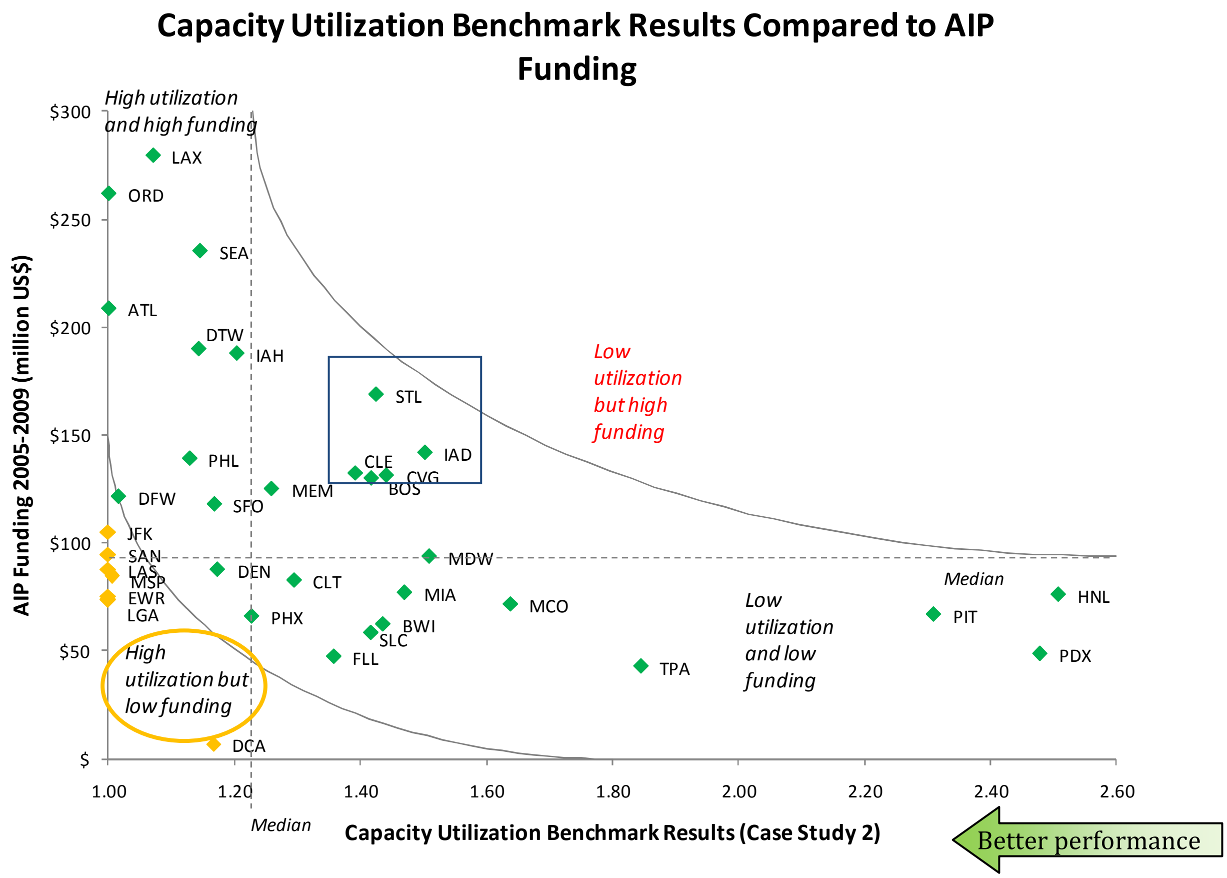

Figure 5.3 - Capacity utilization benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009)

The capacity utilization benchmark results indicate the degree to which airports are using available capacity to maximize the level of air service and the volume of passengers. The results present several insights for policymakers, as indicated by the four areas marked in Figure 5.3:

1. **High utilization and high funding:** This group of airports is operating at high levels of utilization and receives high levels of improvement funding. This indicates that this funding is aligned with areas of high need.
2. **High utilization but low funding**: This group of airports, including EWR and LGA, exhibit high levels of capacity utilization but receive low levels of improvement funding. This should suggest to policymakers a high degree of risk in these airports' ability to accommodate future growth in demand unless further improvements are made. This need for further improvements is subject to the physical possibility of adding further capacity; some airports have physical or other limitations that cause capacity increases to be infeasible. Assuming capacity increases are possible, the airports in this group should be targeted for increased levels of funding.
3. **Low utilization but high funding:** These are airports that receive large amounts of funding in spite of limited capacity utilization. No airports fall in this group, indicating that airport funding is not going

to any airport which does not have a need for improvement funding. However, some airports – including STL, CLE, IAD, and CVG – fall close to this category and receive high levels of funding but according to the benchmark results they do not exhibit a high level of capacity utilization. At all four of the airports listed, improvement funding was allocated for constructing new runways ranging from 44% to 80% of total improvement funding at the airports, as shown in Figure 5.4. This is in spite of comparatively low capacity utilization at these airports. These appear to be cases where funding has not been allocated in a way most consistent with needs since these funds could potentially have been better spent at other airports.

4. **Low utilization and low funding:** This group of airports – including TPA, PIT, HNL, and PDX – has no need for capacity improvements since current capacity utilization is low. These appear to be cases where funding is consistent with needs.

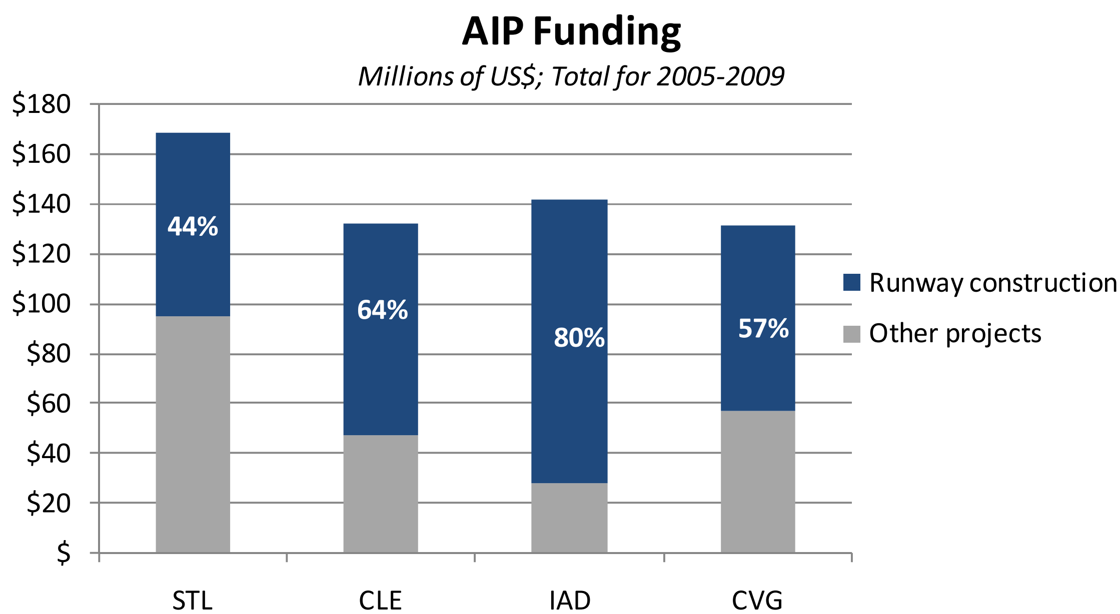

Figure 5.4 - AIP funding directed to runway construction projects

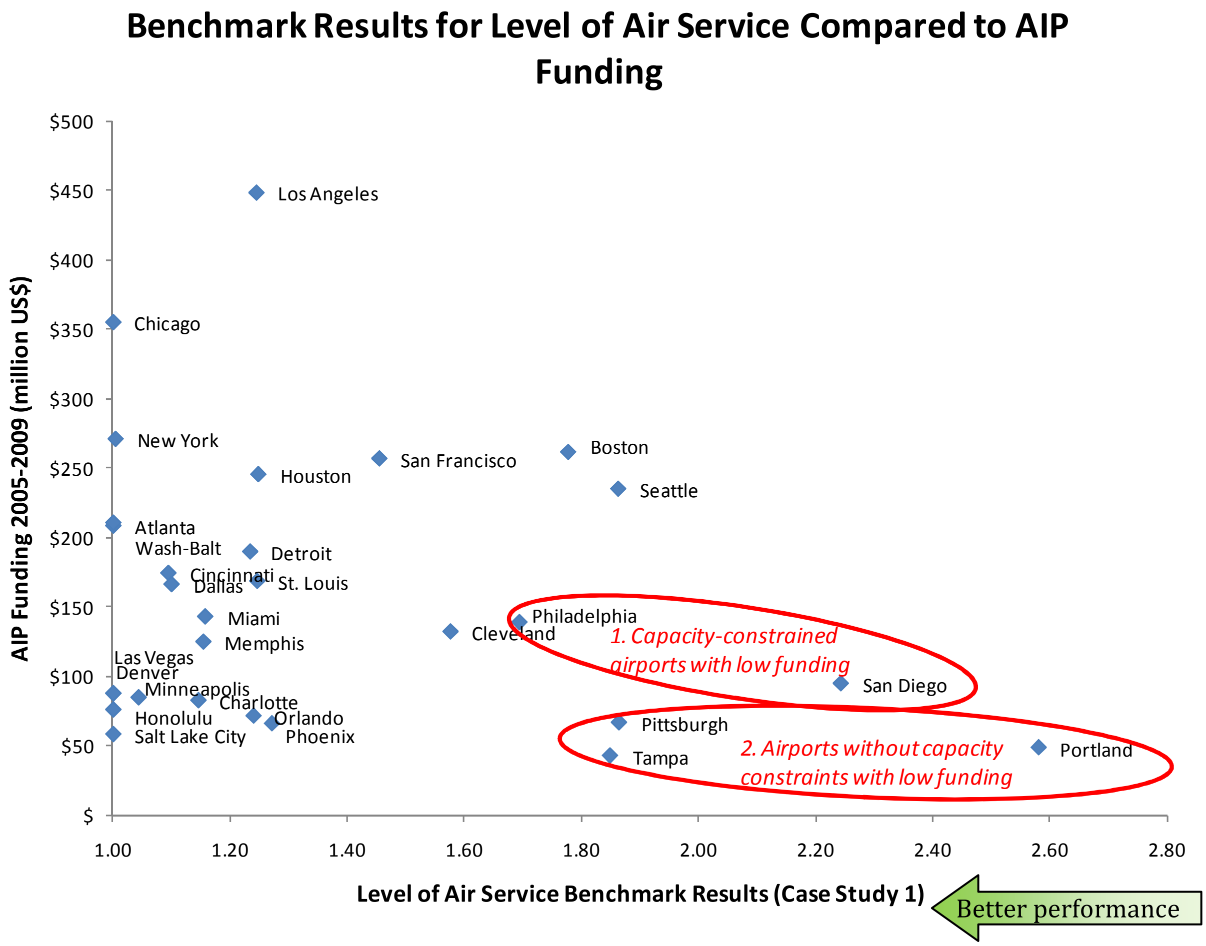

Figure 5.5 - Level of air service benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009)

The level of air service benchmark identifies those regions which have high levels of air service in comparison to the size of their regional economy and population. The metropolitan areas shown in the right hand side of Figure 5.5 are those which have poor levels of air service in comparison to the size of the regional economy. Among those areas, two groups which have particular implications to policymakers can be identified, as indicated in the figure:

1. **Capacity constrained airports with low funding**: As discussed in section 4.2.3.3, Philadelphia and San Diego are among the areas with poor levels of air service and capacity limitations. In the figure, they are also indicated as areas with relatively low levels of improvement funding. Two different causes may exist for this: Either there are limited improvements that are physically possible (e.g. no second runway can be added at SAN), or improvement are possible, but have not been financed by the AIP. If capacity improvements are in fact possible, Philadelphia and San Diego are areas to which increased funding should be allocated.
2. **Airports without capacity constraints and low funding**: The figure indicates that Pittsburgh, Tampa, and Portland are among the areas with poor air service that receive limited improvement funding. As supported by the analysis in section 4.2.3.3, this appears to be a case where funding is consistent with needs. There is no evidence that air services would improve to these areas with further funding since the airports already have excess capacity.

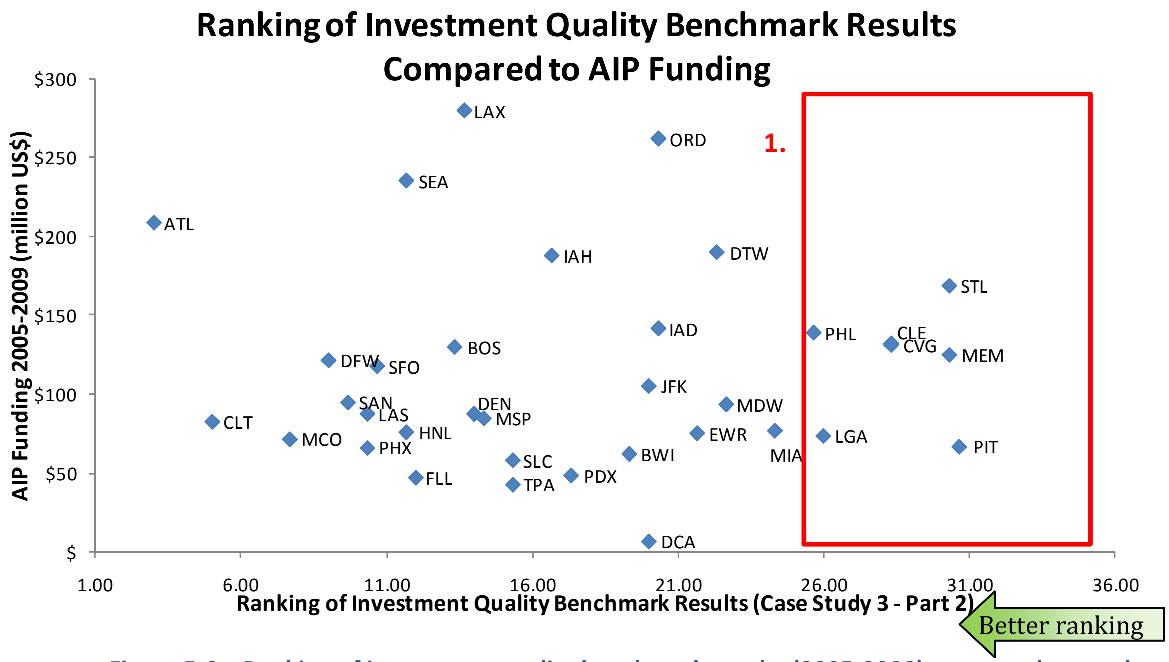

Figure 5.6 – Ranking of investment quality benchmark results (2005-2008) compared to total AIP spending by region (fiscal years 2005-2009)

The investment quality benchmark results are shown in conjunction with the levels of improvement funding in Figure 5.6. This benchmark indicates the degree to which airports would be attractive to investors based on their characteristics of air service, financial factors, and the growth profile for the region.

These results provide one important implication for policymakers: There is a group of airports (in the right-hand side of the figure) which are not attractive targets for investment by private sources of improvement capital. They are not attractive candidates because of poor performance in some or all of the categories of factors considered in the benchmark; for instance, they might be in a region where growth projections are poor, or they might have low levels of non-aeronautical revenues and a poor debt service coverage ratio.

This group of airports is likely to have more difficulty in obtaining approval for funding from private sources, and if such approval is granted, they are likely to be paying higher interest rates than other airports. If improvements are needed at any of these airports – and PHL appears to be an airport with improvement needs, as discussed in the earlier portion of this section – then they may be in greater need of federal funding since the availability of other sources of capital is more limited.

5.4.4 Motivations of Airport Management

The stakeholder model shows policymakers that airports must be viewed in the context of the stakeholders they serve. The stakeholder shows that airport management is not primarily motivated by the airport's role in the overall NAS but rather by its interactions with all of the airport's local stakeholders. Only by understanding those local relationships can the policymaker get insight into the incentives and motivations for airport management.

The stakeholder model shows that the airport plays an integral role to the region it serves, functioning in a utility-like role rather than as a revenue source. The airport serves as an economic engine through on-airport economic activities as well as by enabling business transactions for local companies, which generates jobs for local residents. This is an important motivation for the local government that owns the airport.

Airport management is not in control of many of the factors necessary for providing a full air service, and instead collaborates with service providers. These include airlines, ground transportation providers, and concessionaires. This results in airport management having limited control over some factors which contribute toward the negative perception of airports; notably this includes delay levels, which are a result of airline over-scheduling practices. This may also results in airport management not maintaining as strong a focus on those factors which it determines it cannot control.

Although U.S. airports cannot generate profits for their owners, the stakeholder model shows that strong financial performance is an important concern for airport management for three reasons:

1. **Generating a surplus for infrastructure improvements**: By generating a financial surplus, airports can self-finance a portion of improvement projects at the airport.
2. **Maintaining strong credit ratings**: By generating high revenues and minimizing costs, airport management can contribute toward keeping a strong credit rating, which will make access to private improvement capital easier and less costly.
3. **Attracting air service**: By minimizing costs and maximizing non-aeronautical revenues, airports can keep costs to air carriers low. This is a consideration for air carriers when they determine which region to serve.

From this, policymakers should take away the understanding that rather than viewing themselves primarily as nodes in an overall air transportation system, airport management is more likely to make decisions motivated by more immediate concerns that pertain to the region it serves and the service providers with which it collaborates.

5.4.5 Airports' Roles in their Regions

With the view described in section 5.4.5 of airports as integral parts of the regions in which they exist, stakeholders should study the performance of airports in meeting the needs of their regions to understand the motivations of individual airports. From the benchmark of the level of regional air service described in section 4.2, policymakers can find several groups of metropolitan areas, each with unique motivations:

- **Metropolitan areas with high levels of air service and medium-size or large regional economies and populations**: These areas include Atlanta, Chicago, Washington-Baltimore, and New York City. These regions combine medium-size or large regional economies and population with major airline hub service. For these regions, an important concern is to ensure that sufficient capacity exists for serving future growth in demand; if that capacity is not brought about in spite of continued growth in the regional economy and population, it may in the long term cause an impediment to growth. No motivations appear to exist in the stakeholder model for airport management to in any way restrict the usage of airport capacity for connecting traffic in favor of using that capacity for O&D travel.

- **Metropolitan areas with high levels of air service in spite of comparatively small regional economy and population**: The areas in this group include Honolulu, Las Vegas, Denver, Salt Lake City, and Cincinnati. These areas attract high levels of air service either through their roles as important leisure markets or through high levels of connecting hub service. In the case of the latter, a strong motivation exists for airport management to defend the airport's hub status, both to ensure that the region continues to receive strong levels of air service and also to ensure that the airport continues to generate revenues from those connecting passengers. If these airport had to rely on O&D traffic alone, passenger and aircraft movements would be far lower.
- **Under-served metropolitan areas**: For some of the under-served metropolitan areas, which include Tampa, Seattle, Pittsburgh, and Portland, a key focus for airport management will be to attract increased levels of air service to ensure that the region is better served to make better use of the existing infrastructure. Of particular interest for all of these areas would be to attract new or increased levels of hub service. For some under-served metropolitan areas such

as San Diego and Philadelphia, the focus will be to add infrastructure capacity to permit more air service, as described in section 5.4.3.

5.4.6 Conflicting Objectives among Airport Stakeholders

The stakeholder model provides several cases where conflicting objectives exist, sometimes between groups of stakeholders and other times within the same stakeholder group. Understanding these relationships is important for policymakers in determining the likely actions and reactions to change initiatives by stakeholder groups and airport management. These include:

- **The growth of airports**: Opposing interests exist in terms of the growth of airports, as indicated by the positively and negatively reinforcing loops shown in section 2.1.3.4.3. Many groups, including local businesses, airport concessionaires, and residents who work at the airport or at organizations that generate business in some way connected to the airport all benefit from increased activity at the airport. In contrast, regional residents affected by noise and emissions are likely to oppose further growth in activity at the airport.

- **Fares/yields**: Airport management must consider the needs of both passengers and airlines, yet the former group wants to see low fares while the latter has an interest in maximizing yields. Airport management plays a role in this situation and can choose whether or not to actively pursue increased competition at the airport through increased air service from other carriers.
- **Ground transportation**: For the airport, ground transportation provides a source of revenue through parking fees, taxi fees, etc. Meanwhile, maintaining environmental sustainability is also a goal for airport management, and maximizing the volume of travelers accessing the airport through public transit helps achieve that goal. This is an example of where the financial incentives are in conflict with the environmental sustainability objectives.

5.5 Future Work

This dissertation presents to several opportunities for continued work. The opportunities are both of a methodological and application nature.

The methodological future work would further improve the quality of the benchmarking methodology presented in the dissertation. Two such methodological improvements are:

1. **DEA validity**: Development of a measure of validity of DEA results. As noted in (Morrison 2009), DEA lacks gauges of the explanatory power and significance of a model that the R^2 and confidence levels provide for a regression analysis. This places the impetus on making the correct selections about which metrics to include in the study and which DEA model to use. Although the methodology presented in this dissertation improves the reliability of those selections, a measure of the validity of the results would further strengthen audiences' confidence in them.
2. **Elicitation of stakeholder preferences**: Expansion of a method for eliciting stakeholder preferences and incorporating them into the weights applied in the benchmarking analysis. This would build on the work of (Alodhaibi et al. 2010) which was described in section 2.2.2.5. Such a method should include two components:
    i. A description of practical approaches (e.g. survey methods, interview techniques) for eliciting subject-matter expert and

stakeholder preferences among the different performance metrics considered.

ii. A description of how to integrate the preference weights into the application of DEA.

3. **Stakeholder analysis for other geographies**: Expansion of the stakeholder analysis to other, non-U.S. geographies. Since airport ownership forms, regulations, and stakeholder relationships differ between countries, the stakeholder model presented in this dissertation is not wholly applicable to analysis of airports in other geographies. However, an analysis similar to what was conducted in this analysis would generate other airport stakeholder models, which would enable stakeholder-based benchmarking of airports in other geographies.

The opportunities for further applications of the benchmarking methodology are numerous. Two opportunities include:

1. **Benchmarking of airlines' total cost per operation or passenger at U.S. airports**: Although the direct costs paid to the airport in the form of aeronautical charges are an important cost for air carriers, other sources of air carrier costs at an airport also exists, such as the

costs incurred due to delays at the airport. What this total cost is and how it compares between airports is data which has been sought after for some time by airlines and airport management (Hazel 2010). A comprehensive benchmark of total costs to air carriers could support airline decision-making about which airports to serve and could provide objective data for use in negotiations. From the airports side, it could be a competitive tool for those airports that exhibit low total costs.

2. **Benchmarking of airports' environmental and noise performance**: No benchmark exists of the environmental and noise performance of airports exists. Environmental and noise impacts result from many sources, such as aircraft, ground transportation, and on-airport ground vehicles. A comparative measure of airport performance in this regard is of interest to local residents and airport management since it would help adversely affected residents in areas where noise and emissions are high in negotiations with airport decision-makers, and it would help airports whose noise and emissions profiles are good to objectively support their claims.

Appendix A: A Framework and Heuristics for DEA Model Selection in Airport Benchmarking

Section 2.2.2.6 shows that a variety of different DEA models exist and the review of airport benchmarks in section 2.2.5 shows that studies of airport performance have applied several different DEA models. The existence of different model variations, the lack of consistency in their application on the same problem domain, and the impact of model selection on benchmark results (Schaar & Sherry 2008), point to the need for an analysis of the DEA methodologies for airport benchmarking.

(Kleine 2004) provides a general model framework for categorizing DEA approaches and (Gattoufi et al. 2004) provides a broad DEA taxonomy. In this section, these two approaches are combined with the overview of DEA models from section 2.2.2.6 to create a framework and heuristics for DEA model selection.

This section is organized as follows: First, existing frameworks for categorization of DEA models are analyzed. Next, a new, extended framework for the selection of DEA models is presented. In the next section, heuristics for making choices in the DEA framework when modeling airport performance is presented. Finally, computer implementations of DEA models are presented in the last section.

A.1 Existing Frameworks for Analysis of DEA Models

A generic framework for analyzing the attributes of different DEA models is proposed in (Kleine 2004) and is shown in Table A.1. Kleine separates his framework into an analysis of 1) the scalarizing function, which is the function by which each DMU's DEA "score" is a computed (the score being a scalar value, end hence the term "scalarizing"), and 2) the technology, which is the set of underlying characteristics of the production technology (i.e. the method of converting inputs to outputs) being studied. Kleine categorizes a number of different DEA methodologies (some of which fall outside the scope of the models described in section 2.2.2.6) according to these characteristics. Although Kleine's model classifications list the Additive model as having simple weights, they should in fact be classified as "specific" based on the model description in section 2.2.2.6.3.

Table A.1 - A DEA classification framework (Kleine 2004)

DEA model	Scalarizing Function: Aggregation	Weights	Orientation	Technology
CCR Model	(ε)-maximin	specific	yes	CRS
BCC Model	ε-maximin	specific	yes	VRS
ST Model	ε-maximin	specific	yes	VRS
FDH Model	ε-maximin	specific	yes	FDH
Non-Convex Model	ε-maximin	specific	yes	$NIRS^{FDH}$
Graph-Farrell Meas.	ε-maximin	specific	no	…
Additive Model	additive	simple	no	VRS
MIP Model	additive	specific	no	VRS
RA Model	additive	range-adj.	no	VRS
Russell Measure	additive	specific	yes	…
RA-Graph Model	ε-maximin	range-adj.	no	…
Euclidean Measure	Euclidean	…	…	…
Indivisible	…	…	…	CRS_N
Activity-Limited Model	…	…	…	IBRS

The scalarizing function has three attributes:

- **Aggregation**: This is the means by which the individual components are combined into a scalar value in the objective function. “Additive” indicates that all components are added up while “maximin” is reflective of the ability to remove the impact of some inputs/outputs by assigning weights of zero. “ε-maximin“ reflects the requirement of some models that all weights be non-zero. “Euclidean” refers to the use of a Euclidean norm in the aggregating function (Saneifard et al. 2007).
- **Weights**: This determines how the weights for each parameter are determined. “Simple” refers to the same, standard weight being used for every parameter. “Range-adjusted” indicates that a different weight is used for each parameter, but that the same set of weights are used for each DMU. “Specific” indicates that a different weight is used for each parameter, and that unique weights are determined in the calculation of each DMU’s score.
- **Orientation**: This indicates whether or not models distinguish between input and output orientation, as described further later in this section.

The technology element in the framework has a several possible values. "Technology" refers to the underlying assumptions in the model about how firms (DMUs) are able to convert inputs to outputs. Kleine's paper includes a hierarchical model of different technologies, with CRS at the top of the hierarchy and all other technologies as children. The elements of the framework included in Table A.1 are a subset of all the technologies described in Kleine's paper. The elements in the table translate as follows:

- **CRS:** Constant Returns to Scale, as described in section 2.2.2.6.2.
- **VRS:** Variable Returns to Scale, as described in section 2.2.2.6.2.
- **FDH**: Free Disposal Hull, as described in section 2.2.2.6.5.
- **$NIRS^{FDH}$**: Non-Increasing Returns to Scale in combination with FDH. NIRS is a sub-type of VRS.
- **CRS_N**: Constant Returns to Scale in a model which includes inputs and/or outputs that are indivisible and impose integrality constraints.
- **IBRS**: Individual-Bounded Returns to Scale. The feasible region for the production possibility set do not follow a set of general rules and are instead described by problem-specific rules.

A taxonomy for DEA modeling is provided in (Gattoufi et al. 2004). The taxonomy includes not only elements related to the choice of DEA model but also elements that relate to an applied DEA study more broadly (e.g. the characteristics of the data being used). The taxonomy is detailed in its structure but it does not include any definitions of its elements. The lack of definitions limits the use of the taxonomy. Some of the elements of the taxonomy that are not covered in the framework proposed by Kleine include:

- **Characteristics of the data**: The source of the data (e.g. simulated vs. real data), the domain of the data (e.g. the industry from which it stems), and the level of imprecision in the data.
- **Deterministic or stochastic frontier:** The vast majority of DEA analyses are deterministic in nature, although some efforts have been made to introduce stochasticity to DEA (Ray 2004, p. 307).
- **Time horizon:** Does the analysis encompass a single or a multiple time periods, as discussed in section 2.2.2.6.8?
- **Sensitivity analysis:** Are any tests conducted to test the sensitivity of the study results to factors such as the DEA model choice and sample spread.

A.2 Extending a Framework for Selection of DEA Models

The purpose of this analysis is to assemble a framework for making DEA modeling choices. Since section 2.2.5 shows that airport performance studies in the majority of cases use DEA modeling, the scope of the framework is limited to DEA analyses only and does not address any of the other modeling scenarios described in sections 2.2.2.2 through 2.2.2.4.

This section presents a framework for making DEA model choices and is extended from the existing taxonomies and frameworks described in section 0.

This framework focuses on the choices specific to the DEA model. While several other steps mentioned in the taxonomy of (Gattoufi et al. 2004) such as determining the characteristics of the data, are important to structuring the DEA analysis, those elements are not treated in this framework since they do not pertain to specifying the DEA model itself but rather relate more broadly to how to conduct a DEA study. Practical implementation guidelines for analysis using DEA are available in (R. G. Dyson et al. 2001), for example.

The process for creating the new framework is described in Figure A.1.

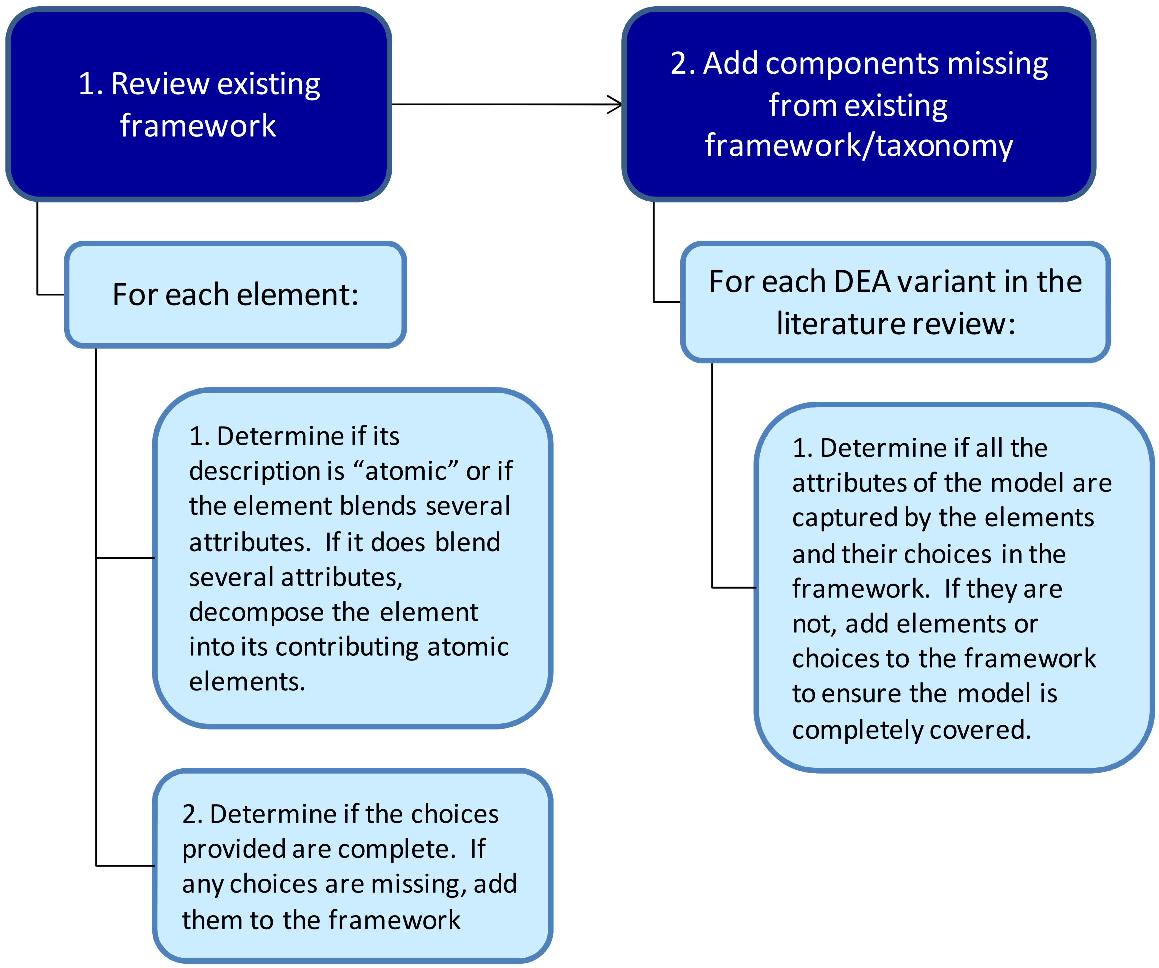

Figure A.1 – The process for creating the new DEA model selection framework

A summary of the new framework is presented in Table A.2 and the subsequent sections address each of the elements in the framework.

The key differences between this new framework and the framework of (Kleine 2004) are that:

1. The technology element has been expanded into several sub-components.
2. A “timespan” element has been added to the framework
3. A “tie-breaking” element has been added to the framework
4. The treatment of negative values has been incorporated into the framework.

Table A.2 - Structure of a DEA model framework for airport benchmarking. Each section in the framework describes an element for which the DEA modeler must make a selection from the choices presented.

Scalarizing function
Aggregation
• ε-maximin
• Maximin
• Additive
• Additive with tolerance for negative values
Weights
• Simple

- Range-adjusted
- Specific

Orientation

- Input
- Output
- None

Technology

Returns to scale

- Constant
- Variable
- Non-increasing
- Non-decreasing
- Individual-bounded

Free Disposal Hull

- Yes
- No

Integer constraints

- All variables integer constrained

• Some variables integer constrained
• No variables integer constrained
Timespan
• Single time period
• Multiple time periods with Malmquist
• Multiple time periods without Malmquist
Tie breaking
• None
• Super-efficiency
• Radii of classification preservation
• Inefficiency frontier

A.2.1 Aggregation

The aggregation method is the means by which the individual components are combined into a scalar value in the objective function. In this framework, unlike the framework proposed in (Kleine 2004), the Euclidean aggregation function is not included as no applied study in any domain applying this aggregation function could be found.

The implications of each type of aggregation function are now reviewed.

A.2.1.1 Maximin

In a maximin-based aggregation function, the model allows for minimization of the impact on the resulting score of any undesirable inputs or outputs, and maximization of the benefit of desirable inputs or outputs. The underlying implications in this modeling choice are that:

- Each DMU is making a choice in allocating attention in order to achieve strong performance on one or more inputs and outputs over other inputs and outputs. The maximin-based aggregation function assumes that this mix of inputs and outputs that the management has chosen is the correct one for the DMU, and that the improvement target for the DMU is one where all the outputs are proportionally increased or all the inputs are proportionally decreased, preserving the relative mix of inputs or outputs.
- Poor performance on some inputs and outputs can be ignored in considering a DMU fully efficient.
- It is acceptable that two DMUs exist on the efficient frontier even though one has worse performance on one parameter than the other

(this is possible through the ignorance of slack in the objective function).

The analyst must make an analysis of the implications of deeming one or several DMUs fully efficient when some inputs and outputs are ignored when selecting the maximin aggregation function.

A.2.1.2 ε-maximin

The ε-maximin function includes the same assumptions as the maximin function, with the addition of the constraint that all parameters being summed have a lower bound of ε. The added constraint implies that poor performance on some parameters cannot be ignored but its impact can be minimized. Using an ε-maximin aggregation function can be one means for the analyst to address the limitations of a maximin function when performance across all parameters must be considered in computing the performance score.

The analyst's choice of value for ε in the DEA implementation may have implications on the results, as discussed in (Ali & Seiford 1993). The authors studied the choice of ε in the CCR and BCC models and provide guidance on how to treat this problem in DEA implementations.

A.2.1.3 Additive

The additive aggregation function implies that all weighted slacks are added up in the objective function. This aggregation function shares the attribute with the ε-maximin function that no parameters can be fully ignored, but the difference between the two functions is that the additive function is a non-radial means of computed a weighted sum of all slacks while the ε-maximin's radial approach represents the proportional distance from the DMU to the frontier.

Different implementations of the additive aggregation function exists, including the original additive function which lacks unit independence (as described in section 2.2.2.6.3) and the SBM/Russell Measure of efficiency which was derived to provide a units-independent version (as described in section 2.2.2.6.4).

A.2.1.4 Additive with Tolerance for Negative Values

The additive aggregation function with tolerance for negative values reflects the same underlying assumptions as the additive aggregation functions in the previous section. The difference in this aggregation function can be characterized as a technical refinement to the model in order to provide the ability for the modeler to handle negative values, rather than a change in the underlying assumptions of the model. The only known implementation of the additive function with tolerance for negative values is the additive model adjusted for negative data described in section 2.2.2.6.11.

A.2.2 Weights

The weights in the DEA model determine the relative importance of the different parameters in the analysis and may also serve to adjust for variations in magnitude between different parameters, depending on the method chosen for determining weights.

A.2.2.1 Simple

Simple weights are the equivalent of two sets of arithmetic means, one for the inputs and one for the outputs, since one common weight is applied to all inputs and one is applied to all outputs.

A.2.2.2 Range-adjusted

Range-adjusted weights are used to adjust for differences in parameter magnitude and are commonly implemented (Kleine 2004, p. 21) as the inverse of the maximum value of each parameter across all DMUs. This method results in a means of combining inputs and outputs of different magnitude that assigns a proportionally even level of importance to each input and output.

This method is appropriate in cases where the analysis requires that each input and output be considered in a proportionally even manner. The underlying assumption is that no choices would be made by the DMUs to emphasize performance in one area over another, and also that, jointly across all DMUs, no preference exists which would warrant a proportionally heavier weighting assigned to one input or output over another.

A.2.2.3 Specific

The specific method of determining weights means that each DMU is assigned its own set of weights for each parameter. The assumption implicit in this method is that, similar to the maximin aggregation function, the DMU has made choices about which inputs and/or outputs to focus on in achieving the best possible performance. To reflect this assumption, the ideal weights for each DMU, subject to the constraints of the DEA model, should be assigned individually for each DMU.

A.2.3 Orientation

The orientation parameter refers to whether the model is input/output-oriented or not, as originally introduced in section 2.2.2.6.2. Models that lack orientation are those that produce the same efficiency score whether inputs are minimized or outputs are maximized. In contrast, the objective function for input-oriented models looks to minimize inputs while keeping outputs constant and output-oriented objective functions look to maximize outputs while keeping inputs constant.

Whether or not the model is orientation-specific or not is dictated by the type of returns to scale specified in the model choice. For instance, CRS models by definition lack orientation.

The studies in Table 2.8 represent a mix of both oriented and non-oriented models: Half of the studies use an oriented model, 17% use models without orientation, and the remaining third of studies use a mix of both input and output-oriented models. For those studies that use oriented models, there is a roughly even mix of input and output-oriented model. Several studies explain why input or output-oriented models were used, and a common motivation is to choosing the orientation depending on what is considered most controllable by airport management. For example, minimizing inputs is considered within the control of management, and as a result the model chosen is input-oriented (Abbott & Wu 2002, p. 252). Meanwhile other studies consider output maximization more controllable by management since once the inputs are created, they cannot be changed, and as a result an output-oriented model is used; "once an airport has invested in the building of new runways or new terminals, it is difficult for managers to disinvest to save costs” (Martín & Román 2001, pp. 152-153).

A.2.4 Returns to Scale

The returns to scale element of the taxonomy indicates the functioning of the underlying technology in growing output volumes as inputs increase.

The CRS assumption reflects the reality that efficiency does not improve or worsen as the scale of operations increase but rather that the returns are proportional to any growth or shrinking in the inputs. In contrast, VRS implies that increasing or decreasing returns to scale may exist.

Non-Increasing Returns to Scale (NIRS) means that increasing the scale of operations yields at best constant returns, but may also yield decreasing returns to scale, and the Non-Decreasing Returns to Scale (NDRS) assumption implies that increasing the scale of operation results in constant or increasing returns. These are both sub-sets of VRS.

Lastly, the Individual-Bounded Returns to Scale (IBRS) assumption means that no standard rules can be described for scale returns, and that instead individual bounds for the production possibility set must be described. This is the result of only certain levels of activity being feasible in the area that is being modeled. A literature search across all modeling domains reveals no cases of this assumption having been used in past models.

A.2.5 Free Disposal Hull

As discussed in section 2.2.2.6.5, Free Disposal Hull assumes that efficiency of airports can only be measured relative to observed sets of inputs and outputs rather than relative to linear combinations of points on the efficiency frontier, as assumed by many standard DEA models such as CCR and BCC.

The motivation for FDH has been described by (Tulkens 1993): "…the identification of a set of dominating observations, by showing actually implemented production plans that are clearly more efficient, gives to the inefficiency scores a credibility that they usually lack when reference is only made to an abstract frontier." This motivation points not to any fundamental difference in the production technology as compared to other DEA methodologies, but rather to the advantage of achieving stronger acceptance of the benchmark results when comparisons are made by providing a comparison to only one peer DMU.

A.2.6 Integer Constraints

The question of integer constraints on some inputs and/or outputs is embedded in the framework of Kleine but is not broken out as a separate parameter in the framework. However, in (Zhu & Cook 2007, pp. 271-273) Lozano and Villa point out the importance of integer constraint considerations for some inputs and/or outputs for parameters of relatively small magnitudes where fractional values are infeasible (e.g. the number of workers, number of machines, etc.). The authors conclude that in DEA analyses where all inputs and outputs have high magnitudes, integer constraints are not an important consideration even when those parameters are indivisible since rounding ex post introduces a very small error. However, for airports inputs such as the number of runways meet the criteria of indivisibility and small magnitudes. As a result, the presently proposed framework breaks out integrality as a separate consideration.

In the work of Lozano and Villa (Zhu & Cook 2007, p. 272) which was published in 2007, the authors point out that in spite of the possibility of indivisibility of inputs and outputs may occur frequently, no other authors have addressed this issue.

A.2.7 Time Horizon

In the event that the analysis contains repeated observations over time about the same set of DMUs, the analyst must determine how to treat these observations. As discussed in section 2.2.2.6.8, three modeling options exist:

1. Combining all measures into a single analysis, allowing a DMU to be benchmarked against its own performance (as well as that of all other DMUs) at different time periods.
2. Computing scores in separate analyses, comparing each DMU only against its other peers for the same time period.
3. Computing a Malmquist index version of the DMU model being considered, accounting for any changes in the underlying production technology.

A.2.8 Tie-Breaking

In some modeling circumstances, there is a need for all DMUs to be fully ranked. The general DEA models create an efficiency frontier which permits several DMUs to be present on the frontier at the same time and the models also make it possible for several DMUs to take on the same efficiency score. To break this tie, three different approaches have been proposed, as described below. It should be noted that none of the approaches guarantee that all ties be broken; rather they strongly reduce the probability that ties exist.

1. **Super efficiency**: The super efficiency approach, as described in section 2.2.2.6.6, has two different variations:
    a. Removing the DMU for which the efficiency score is being computed from the constraints section, making it possible for that DMU to achieve a score higher than 1.0 in the input oriented model and lower than 1.0 in the output oriented model.
    b. Creating an artificial DMU which takes on the highest recorded value for each output and the lowest recorded value for each input.
2. **Radius of Classification Preservation**: The RCP approach is described in section 2.2.2.6.9 and is only applied to the DMUs that are on the efficiency frontier. It tests the sensitivity of the DMUs' classification as efficient by testing the degree to which the inputs and

outputs can be changed before the classification changes from efficient to inefficient.

3. **Inefficiency frontier:** The inefficiency approach is described in section 2.2.2.6.10 and involves creating a frontier of the fully inefficient airports which are not "negatively dominated" by any other airport, and measuring the distance from this frontier for all other airports. This distance from the inefficiency frontier is used as a tie breaking mechanism.

A.3 Developing Heuristics for DEA Modeling of Airport Performance

This section presents heuristics, or "decision rules", for making selections in the DEA model framework from section A.2 when modeling airport performance. These heuristics were developed by analyzing existing research about the characteristics of airport performance as it pertains to the elements of the DEA framework with the objective of providing decision guidance for what the appropriate selections in the framework are, based on the aspect of airport performance being modeled.

Each of the following subsections addresses one aspect of the DEA framework, and follows the same order as the framework. The last subsection provides a summary of the framework and mapping to DEA models.

A.3.1 Aggregation

Among the studies in Table 2.8, all but one apply models which use maximin or ε-maximin aggregation. The exception is the GTR model whose aggregation function is additive. In the two models which make up the original DEA formulations, CCR and BCC, the aggregation function is maximin or ε-maximin, and it is conceivable that the reason for the prevalence of maximin or ε-maximin aggregation functions is due to some authors of studies in Table 2.8 choosing CCR and BCC by "default".

The selection of aggregation function in airport DEA models should be based on the underlying characteristics of the aspects of airport performance that are being studied.

First, (R. G. Dyson et al. 2001, p. 253) advocate that all parameters should be included in the objective function calculation since there "has to be an agreement on which factors to include in the DEA assessment, and that agreement implies that the factors are important and should be taken account of". This reasoning results in only ε-maximin or additive aggregation function being appropriate choices since the maximin function allows for some inputs and outputs to be fully ignored.

When choosing between the ε-maximin and the additive function, three factors should be of concern: The first is: How should the results be expressed and interpreted? The ε-maximin function is interpreted as the proportional increase in outputs (or decrease in inputs) in order to reach the efficient frontier, assuming no change in relative mix between outputs (or inputs). The additive function lacks any interpretation along these lines, making the ε-maximin function the "default choice" for modeling as it maintains the underlying assumption that management at the DMU has made mix decisions that are optimal of the DMU's particular context.

The second question is: Is the ignorance of slack acceptable in the objective function? If the answer is yes, then the ε-maximin function remains a viable choice; if not, then the choice should be the additive function. If an additive function is to be used, the units-independent version implemented in the SBM/Russell Measure of efficiency should be used (as described in section 2.2.2.6.4).

Finally, the third concern is: Does any parameter take on negative values. If that is the case, then the only model that can be used is the additive function that is implemented in the additive model adjusted for negative data (as described in section 2.2.2.6.11).

A.3.2 Weights

All of the studies of airport performance listed in Table 2.8 use specific weights for each parameter and each DMU. This reflects the assumption that parameters must be scaled based on their relative magnitude, but also the fact that not all parameters carry the same proportional importance, and that those relative levels of importance may vary for each DMU.

Simple weights are in most cases not an appropriate method to use in comparative airport studies since parameters generally do have differing levels of magnitude and may have different relative levels of importance. Range-adjusted weights can be used but should only be applied in those cases where using the proportional weights that assign equal importance to each parameter can be motivated in the analysis; a reason must exist why all parameters are of equal importance.

If parameters do have differing levels of magnitude and if no motivation can be provided as to why each parameter has proportionally the same level of importance, then specific weights should be used in analyses of airport performance.

A.3.3 Orientation

The choice of returns to scale in the DEA model will determine whether the model is oriented or not. If the returns to scale choice results in an oriented model, then the analyst has to select input or output orientation. The analysis in section A.2.3 indicates that the analyst should determine whether inputs or outputs can be considered controllable by management, and based on that determination, the model orientation should be determined.

A.3.4 Returns to Scale

Half of the studies in Table 2.8 assume VRS, 17% assume CRS, and one third of the studies assume some combination of the two as a result of running multiple different models. Few of the studies discuss why the VRS and/or CRS model was chosen.

The issue of CRS and VRS is one of the most studied characteristics of production frontiers in economics (Rajiv D. Banker 1996, p. 148). (Ray 2004, p. 46) point out that "…it is unlikely that CRS will hold globally in many realistic cases. As a result, the CCR-DEA model [which assumes CRS] should not be applied in a wide variety of situations". As a counterpoint, "…if the VRS model is used, where there are no inherent scale effects, small and large units will tend to be over-rated in the efficiency assessment" (R. G. Dyson et al. 2001, p. 248).

One airport study points out that in selecting CRS or VRS "it is especially important to have some idea about the hypothetical returns to scale that exist in the industry" (Martín & Román 2001, p. 152). These authors point to the importance of ensuring that the fundamental modeling assumptions reflect the reality that is being modeled.

In selecting the types of returns to scale, the analyst must take care to reflect the real-world conditions that are being modeled, as suggested by (Martín & Román 2001, p. 152). This may mean gaining a qualitative understanding of the domain being modeled to determine what types of scale returns exist. (R. G. Dyson et al. 2001, p. 248) advocates testing the data for scale effects and using the VRS model only when scale effects can be demonstrated. (Rajiv D. Banker 1996, pp. 148-151) proposes a general method for testing for the existence of VRS.

In an extensive study of returns to scale across international airports (Martin & Voltes-Dorta 2008), significant evidence of increasing returns to scale was found.

The study considered labor costs, capital costs, and material costs as inputs; and air traffic movements, passengers, and cargo volumes (passengers and cargo were combined into the measure of Workload Units [WLUs], which are defined as either one passenger or 100 kg of cargo) as outputs. The study included data from 41 airports across Europe, North America, Asia, and Australia, and studied performance from 1991 to 2005. The study authors performed comprehensive data preprocessing to account for differences in operating models and data reporting methods. An analogous study cited in (Morrison 2009) of 36 airports across different geographies from 1993-2000 using similar inputs and outputs also confirmed the findings of increasing returns to scale.

These results indicate that in studies that involve resource inputs like labor and capital resources and outputs like ATMs and WLUs, the choice of returns of scale should be VRS or NDRS.

A.3.5 Free Disposal Hull

Among the airport studies in Table 2.8, none have applied an FDH model. The decision about whether or not to use of FDH in airport benchmarking should be based on how the benchmark results are expected to be used. As suggested by (Tulkens 1993), in cases where acceptance of the benchmark results will be strengthened by showing comparisons to only one other airport, the analyst should consider using the FDH methodology.

A.3.6 Integer Constraints

In spite of the frequent use of inputs such as the number of runways, no studies in Table 2.8 take into account these integer constraints.

The analyst of airport performance should consider which inputs and outputs are indivisible and are low in magnitude, and should apply integer constraints to those parameters. No guidance exists in the literature for determining what is to be considered "low magnitude"; however, if in doubt, the analyst should err on the side of caution and apply integer constraints. Examples from past airport studies where integer constraints should be considered include:

- Runways
- Gates
- Baggage conveyor belts

A.3.7 Time Horizons

Eight of the 12 airport studies in Table 2.8 that are DEA-based use observations from multiple time periods, and among those studies, only one applies a Malmquist index. Among the remaining seven studies, six conduct separate analyses for each year and one pools all observations into a single analysis.

Determining whether to compute a Malmquist index depends on whether or not any changes in the underlying technology can be expected over time. For instance, in the case of an airport benchmark that considers the runway to be a resource input and the number of aircraft movements to be an output, the analyst must consider the question whether the introduction of new equipment or procedures during the time period being analyzed would alter the feasible processing rate. If such changes are expected, then a Malmquist index should be computed. In contrast, if the analyst determines that no relevant changes occurred during the analysis time period, then either the first or the second option described in the previous section would be the most appropriate.

Evidence of the existence of technological change in airport operations over time were found in the same study that also identified the existence of variable returns to scale (Martin & Voltes-Dorta 2008). The study reviewed data from 1991 to 2005 and used labor costs, capital costs, and material costs as inputs; and ATMs and WLUs as outputs. The existence of technology changes over time indicates that studies which use these types of inputs and outputs across multiple time periods should compute a Malmquist index.

A practical drawback of computing separate analyses for each year (whether using a Malmquist index or not) rather than a single, combined analysis is that the number of observations in each model run is lower by a factor of t when a separate analysis is done for each year, where t is the number of time periods being analyzed. In cases where the ratio of the number of observations to the number of parameters being considered is low, this can be a limitation to the usefulness of the results.

A.3.8 Tie-Breaking

The tie-breaking function serves to ensure that the results are as fully ranked as possible. Approaches to fully ranking airports were used in (Bazargan & Vasigh 2003) and in (Adler & Berechman 2001).

Each of the approaches to tie-breaking described in section 2.2.2.6 are more complex formulations of the general DEA models, suggesting that their use is moves the model further away from the original interpretation of DEA as formulating a frontier of efficient DMUs and measuring the distance to that frontier. Accordingly, tie-breaking approaches should not be used in airport DEA modeling unless a reason that all airports must be fully ranked exists.

If such a reason exists, the modeler should select the tie-breaking method which has a clear intuitive interpretation; as with the approach to FDH discussed in section A.3.5, this provides results that are more likely to be accepted by the benchmark audience. No general guidance can be provided as to which of the approaches to tie-breaking provides the best intuitive interpretation.

A.3.9 Summary of Airport DEA Model Selection Heuristics

A summary of the heuristics is included in Table A.3, and the heuristics are translated into specific modeling choices in Table A.4.

Table A.3 - Airport DEA Framework and Heuristics

Scalarizing function
Aggregation
Use either ε-maximin or additive. If the ignorance of slacks in the efficiency score is acceptable, then ε -maximin is the choice that reflects management's choices about the mix of inputs and/or outputs. Otherwise, use the additive function. In addition, if any parameters take on negative values then the additive function implemented in the additive model adjusted for negative data must be used.
Weights

Use specific weights unless evidence exists that range-adjusted weights are more appropriate.

Orientation

If the model requires orientation, then choose orientation to reflect which parameters are controllable by management.

Technology

Returns to scale

If modeling some version of labor and capital resources as inputs and passengers and aircraft movements as outputs, then use VRS. Otherwise, study the parameters to determine if VRS or CRS exist.

Free Disposal Hull

Unless compelling evidence that study results will be better accepted if only observed values are used for peer comparisons, do not use FDH.

Integer constraints

Use integer constraints for inputs and outputs with low magnitudes, such as runways.

Timespan

If modeling some version of labor and capital resources as inputs and passengers and aircraft movements as outputs over multiple time periods, then use a Malmquist index. For other domains, review if technology changes over time have occurred.

Tie breaking

If the study requires that all airports be fully ranked, use the tie-breaking function that provides the best intuitive interpretation; otherwise do not use a tie-breaking function.

Table A.4 - Translation of heuristics to specific model choices

Element	Choice	Translation in modeling
Aggregation	ε-maximin	Use CCR or BCC with minimum bounds on weights, as described in sections 2.2.2.6.1 and 2.2.2.6.2.
	Additive	Use SBM/Russell measure, since these provide units invariant modeling options, as described in section 2.2.2.6.4.
	Additive with tolerance for negative data	Use the adaptation of the SBM/Russell measure model with tolerance for negative data, as described in section 2.2.2.6.11.
Weights	Specific weights	Use original model as specified.
Orientation	Input/output	If using an oriented model such as BCC, choose the input or output oriented version as appropriate.

Element	Choice	Translation in modeling
Technology	CRS	If the aggregation function is ε-maximin, then choose CCR, as described in section 2.2.2.6.1. If using some other model, ensure that no convexity constraint such as eλ=1 is present in the model.
	VRS	If the aggregation function is ε-maximin, then choose BCC, as described in section 2.2.2.6.2. If using some other model, ensure that a convexity constraint such as eλ=1 is present in the model.
Free Disposal Hull	Use FDH	Use the FDH implementation as described in section 2.2.2.6.5.
	No use of FDH	Use original model as specified.
Integer constraints	Some/all variables integer constraints	Use the implementation as described in section 2.2.2.6.12.

Element	Choice	Translation in modeling
	No integer constraints	Use original model as specified.
Timespan	Use Malmquist index	Use Malmquist index implementation as described in section 2.2.2.6.8.
	No Malmquist index	Use original model as specified.
Tie-breaking	Use tie-breaking	Use one of the implementations as described in sections 2.2.2.6.6, 2.2.2.6.9, or 2.2.2.6.10.
	No tie-breaking	Use original model as specified.

Appendix B: Matlab DEA Code

B.1 Code for Single DEA Run

```
function DEA(dataFile, numInputs, numOutputs, integerParams, DEAmodel, orientation,
    weightlbs, varargin)

% Inputs:

%    - An .xls data file with the inputs listed first in columns and then then the
    outputs.  The first row is the header with column names.  The DMU names are in
    column A

%    - The number of inputs

%    - The number of outputs

%    - An integer matrix (numInputs+numOutputs,1) with 1s marking those
    inputs/outputs that are integer, and 0s otherwise.

%    Integer parameters are currently only supported for CCR and BCC

%    - The DEA model.  Current choices are (case sensitive): CCR, BCC, SBM, ADDNEG
    (Additive model adjusted for negative data)

%    - The orientation of the DEA model ('input' or 'output')

%    - The lower bounds on the variable weights (usually set to 0 or epsilon)

%    - printMode (optional): 1 (def) - detailed; 2 - simple with headers; 3 - simple
    without headers

%    - outfile (optional): defauls to the input file.txt instead of .xls.  Note that
    this should be the actual file "stream", not the file name

% Outputs: The DEA scores and other parameters

% DEA definitions:

% X: The inputs matrix

% Y: The outputs matrix
```

```
addpath cplexincludes;

%start off with dealing with optional arguments
%print mode
if size(varargin,2) > 0
    printMode = varargin{1};
else
    printMode = 1;
end
% Set output file:
if size(varargin,2) > 1
    outfile = varargin{2};
else
    outfile = fopen(strrep(dataFile,'.xls','.txt'),'w');
end
 %Read in the .xls data file, which has this format:
% DMUs, 'input1title', 'input2title', 'output1title', 'output2title',
% DEA1Name,input1,input2,output1,output2,...
% DEA2Name,input1,input2,output1,output2, ...
% ...
benchmarkData = csv2struct(dataFile);
fns = fieldnames(benchmarkData);
numDMUs = length(benchmarkData.(fns{1}));

X = zeros(numDMUs, numInputs);
for i=2:numInputs + 1
    X(:,i-1) = benchmarkData.(fns{i});
end;

Y = zeros(numDMUs, numOutputs);
for i=numInputs+2:numInputs+1+numOutputs
```

```
    Y(:,i-1-numInputs) = benchmarkData.(fns{i});
end;


% Determine model parameters:
% Test for convexity
if strcmp(DEAmodel,'BCC') || strcmp(DEAmodel,'ADDNEG')
    convexity = true;
else
    convexity = false;
end;
% Structure the LP.  The main constraints section is standard.  What varies by model
      are the objective function and a few other factors.  We'll construct the dual
      problem.
% Size of A: - One row for each input and one for each output.
%             - One extra row for each integer constraint
%             - One row to leave room if convexity constraints apply
%             - One row for the value of theta (or t) in phase 2
%             - One column for the theta (or t)
%             - One column for each DMU
%             - One column for the slack var for each input and output
%             - One column for each integer constraint
% Although the below could be written more concisely, it is spelled
% out here to more clearly match the listings above
A = zeros(numInputs + numOutputs + numInputs + numOutputs + 1 + 1, 1 + numDMUs +
      numInputs + numOutputs + numInputs + numOutputs);
B = zeros(numInputs + numOutputs + numInputs + numOutputs + 1 + 1, 1);
 % The A and F matrices are indexed by integer values.  To be able
% to use proper variable names and indices, we create these arrays
% to keep track of the variable to matrix index mapping:
variables = struct('theta',{},'lambda',{},'xSlack',{},'ySlack',
      {},'xInteger',{},'yInteger',{});
varCounter = 1;
variables(1).theta = varCounter;
```

```
varCounter = 2;
variables.lambda = (varCounter:varCounter + numDMUs - 1);
varCounter = varCounter + numDMUs;
variables.xSlack = (varCounter:varCounter + numInputs - 1);
varCounter = varCounter + numInputs;
variables.ySlack = (varCounter:varCounter + numOutputs - 1);
varCounter = varCounter + numOutputs;
if sum(integerParams(:)) > 0
    % although we technically only need these variables for the
    % integer variables, it's easier to make them for all variables
    % if at least one integer variable is present.
    variables.xInteger = (varCounter:varCounter + numInputs - 1);
    varCounter = varCounter + numInputs;
    variables.yInteger = (varCounter:varCounter + numOutputs - 1);
    varCounter = varCounter + numOutputs;
end;
numVars = size(A,2);
% Create the standard A matrix
A = createBasicA(A, X, Y, numInputs, numOutputs, numDMUs, orientation, DEAmodel,
     variables, integerParams);
% This adds the convexity constraint if necessary; e.g. for BCC
if convexity
    A(end-1,variables.lambda(1):variables.lambda(end)) = 1;
    B(end-1,1) = 1;
end;
%Write the cplex problem
%Determine if this is a min or max problem
if strcmp(DEAmodel, 'SBM')
    %min
    sense=1;
elseif strcmp(DEAmodel, 'ADDNEG')
    %max
```

```
    sense=-1;
elseif (strcmp(DEAmodel, 'BCC') || strcmp(DEAmodel, 'CCR')) &&
     strcmp(orientation,'input')
    %min
    sense=1;
else
    %max
    sense=-1;
end;

%Set constraint types (all ours are equality)
ctype = blanks(size(B,1))';
ctype(1:end) = 'E';
%Set variable bounds
lb = zeros(numVars,1);
ub = inf(numVars,1);
if strcmp(DEAmodel, 'SBM')
    % in the SBM model, t must be positive
    lb(variables.theta) = 0.0000001;
end;
vartype = blanks(numVars)';
vartype(1:end) = 'C';
%Set integrality constraints
if sum(integerParams(:)) > 0
    indexArray = (1:length(integerParams));
    for i=indexArray(ismember(integerParams,1))
        if i <= numInputs
            vartype(variables.xInteger(i)) = 'I';
        else
            vartype(variables.yInteger(i-numInputs)) = 'I';
        end;
        % set constraints that aren't equality
```

```
        if convexity
            if strcmp(orientation,'input')
                ctype(i) = 'G';
            else
                ctype(i) = 'L';
            end;
        end;
    end;
end;

%suppress the full CPLEX output
params.msglev=0;
params.errmsg=0;
%struct for storing results
results = struct('DMUname',{},'efficiency',[],'vars',[],'lambda',[]);
results(numDMUs).vars = [];
%Loop through each DMU and compute its efficiency data
for DMU=1:numDMUs
    [A B] = createDMUSpecificAB(A, B, X, Y, numInputs, numOutputs, numDMUs,
     orientation, DMU, DEAmodel,                                  variables,
     integerParams);
    F = setF(numDMUs, numInputs, numOutputs, DEAmodel, orientation, weightlbs, DMU,
     X, Y, variables, numVars);
    %Run CPLEX    [xopt,opt,status,extra]=cplexmex(sense,[],F,A,B,ctype,lb, ub,
     vartype,[],params,0);
    %Store results
    results(DMU).DMUname = benchmarkData.(fns{1})(DMU);
    results(DMU).efficiency = opt;
    results(DMU).vars = xopt;
    results(DMU).lambda = extra.lambda;
    %Run phase 2, maximizing the slacks
    if (strcmp(DEAmodel,'BCC') || strcmp(DEAmodel,'CCR')) &&
     sum(weightlbs) == 0 && sum(integerParams(:) == 0)
```

```
        Fphase2 = setFphase2(numDMUs, numInputs, numOutputs,
      orientation, variables, numVars);

        A(end,variables.theta) = 1;

        B(end,1) = opt;

        [xopt,opt,status,extra]=cplexmex(sense,[],Fphase2,A,B, ctype,
      lb,ub,vartype,[],params,0);

        %Store phase 2 updated values

        results(DMU).vars = xopt;

        %Clean up phase 2

        A(end,variables.theta) = 0;

        B(end,1) = 0;

    end;

    %re-convert variables for SBM

    if strcmp(DEAmodel,'SBM')

        %divide by t (we are using theta as t)

        results(DMU).vars = results(DMU).vars / results(DMU).vars(variables.theta);

    end;

end;

%DISPLAY THE DATA

%To start, a summary to screen:

fprintf('\nDMU\t\t  Score\n');

fprintf('----------------------------\n');

for DMU=1:numDMUs
     fprintf('%s\t\t%8.4f\n',char(results(DMU).DMUname),results(DMU).efficiency);

end;

fprintf('\n');

if printMode == 1

    %The same summary to file:

    fprintf(outfile, '\nDMU\t\t  Score\n');

    fprintf(outfile, '----------------------------\n');

    for DMU=1:numDMUs

        fprintf(outfile, '%s\t\t%8.4f\n',
      char(results(DMU).DMUname),results(DMU).efficiency);
```

```
    end;
    fprintf(outfile, '\n');
    %Next, a detailed report for each DMU
    fprintf(outfile, '\nDetailed results\n\n');
    for DMU=1:numDMUs
        fprintf(outfile, '------------------\nDMU: %s\n', char(results(DMU).DMUname));
        fprintf(outfile, 'Efficiency: %8.4f\n', results(DMU).efficiency);
        fprintf(outfile, 'Peers:\n');
        for peer=1:numDMUs
            fprintf(outfile, '  %s: %8.4f\n', char(benchmarkData.(fns{1})(peer)),
     results(DMU).vars(variables.lambda(peer)));
        end;
        fprintf(outfile, 'Slacks:\n');
        for var=1:numInputs
            fprintf(outfile, '  s%d-: %18.10f\n', var,
     results(DMU).vars(variables.xSlack(var)));
        end;
        for var=1:numOutputs
            fprintf(outfile, '  s%d+: %18.10f\n', var,
     results(DMU).vars(variables.ySlack(var)));
        end;
        fprintf(outfile, 'Weights:\n');
        for weight=1:numInputs
            fprintf(outfile, '  v%d: %18.10f\n', weight,
     results(DMU).lambda(weight));
        end;
        for weight=1:numOutputs
            fprintf(outfile, '  u%d: %18.10f\n', weight,
     results(DMU).lambda(numInputs + weight));
        end;
    end;
else
    if printMode == 2
        % print row headers for efficiency
```

```
    for DMU=1:numDMUs
        fprintf(outfile, '%s,',char(results(DMU).DMUname));
    end;
    % print row headers for slacks and weights
    for DMU=1:numDMUs
        for var=1:numInputs
            fprintf(outfile, '%s-s%d-,', char(results(DMU).DMUname), var);
        end;
        for var=1:numOutputs
            fprintf(outfile, '%s-s%d+,', char(results(DMU).DMUname), var);
        end;
    end;
    for DMU=1:numDMUs
        for weight=1:numInputs
            fprintf(outfile, '%s-v%d,', char(results(DMU).DMUname), weight);
        end;
        for weight=1:numOutputs
            fprintf(outfile, '%s-u%d,', char(results(DMU).DMUname),weight);
        end;
    end;
    fprintf(outfile, '\n');
end;
%print the data
% print efficiency
for DMU=1:numDMUs
    fprintf(outfile, '%8.4f,', results(DMU).efficiency);
end;
% print slacks
for DMU=1:numDMUs
    for var=1:numInputs
        fprintf(outfile, '%18.10f,', results(DMU).vars(variables.xSlacks(var)));
    end;
```

```
        for var=1:numOutputs
            fprintf(outfile, '%18.10f,', results(DMU).vars(variables.ySlacks(var)));
        end;
    end;
    % print weights
    for DMU=1:numDMUs
        for weight=1:numInputs
            fprintf(outfile, '%18.10f,', results(DMU).lambda(weight));
        end;
        for weight=1:numOutputs
            fprintf(outfile, '%18.10f,', results(DMU).lambda(numInputs + weight));
        end;
    end;
    fprintf(outfile, '\n');
end;
end


function [A B] = createDMUSpecificAB(A, B, X, Y, numInputs, numOutputs, numDMUs,
     orientation, targetDMU,

     DEAmodel, variables, integerParams)
    if strcmp(DEAmodel, 'SBM')
        %assign the t constraint for Y
        A(end,variables.theta) = 1;
        for i=1:numOutputs
            A(end,variables.ySlack(i)) = 1/(numOutputs * Y(targetDMU,i));
        end;
        B(end,1) = 1;
        A(1:numInputs,variables.theta) = X(targetDMU,:);
        A(numInputs+1:numInputs+numOutputs,variables.theta) = -Y(targetDMU,:);
    elseif strcmp(DEAmodel, 'ADDNEG')
            % Assign ya:
            B(numInputs + 1:numInputs + numOutputs,1) = ...
```

```
            Y(targetDMU,:);
        % Assign xa:
        B(1:numInputs,1) = -X(targetDMU,:);
else
    if strcmp(orientation,'input')
        if sum(integerParams(:)) == 0
            % Assign theta:
            A(1:numInputs,variables.theta) = X(targetDMU,:);
            % Assign ya:
            B(numInputs + 1:numInputs + numOutputs,1) = Y(targetDMU,:);
        else
            % We have integer constrained variables
            % Assign theta:
                  A(numInputs+numOutputs+1:numInputs+ numOutputs+ numInputs,
        variables.theta) = X(targetDMU,:);
            % Assign ya:
            B(numInputs + numOutputs + numInputs + 1:numInputs + numOutputs +
 numInputs + numOutputs,1) = Y(targetDMU,:);
        end;
    else
        if sum(integerParams(:)) == 0
            % Assign theta:
                  A(numInputs + 1:numInputs + numOutputs, variables.theta) =
        Y(targetDMU,:);
            % Assign xa:
            B(1:numInputs,1) = X(targetDMU,:);
        else
            % We have integer constrained variables
            % Assign theta:
            A(numInputs + numOutputs + numInputs + 1:numInputs + numOutputs +
 numInputs + numOutputs,variables.theta) = Y(targetDMU,:);
            % Assign xa:
            B(numInputs + numOutputs + 1:numInputs + numOutputs + numInputs,1) =
 X(targetDMU,:);
```

```
            end;
        end;
    end;
end

% This function creates the standard constraints and rhs matrices
function A = createBasicA(A, X, Y, numInputs, numOutputs, numDMUs, orientation,
     DEAmodel,variables, integerParams)
    if strcmp(DEAmodel,'SBM') || strcmp(DEAmodel,'ADDNEG')
        % Assign X:
       A(1:numInputs,variables.lambda(1):variables.lambda(end)) = -X';
        % Assign input slacks
        for i=1:numInputs
            A(i,variables.xSlack(i)) = -1;
        end;
        % Assign Y:
        A(numInputs + 1:numInputs + numOutputs,
     variables.lambda(1):variables.lambda(end)) = Y';
        % Assign output slacks
        for i=1:numOutputs
            A(numInputs + i,variables.ySlack(i)) = -1;
        end;
    else
        % This is CCR or BCC
        if strcmp(orientation, 'input')
            if sum(integerParams(:)) == 0
                % Assign X:
                     A(1:numInputs, variables.lambda(1):variables.lambda(end)) = -X';
                % Assign input slacks
                for i=1:numInputs
                    A(i,variables.xSlack(i)) = -1;
                end;
```

```
            % Assign Y:
            A(numInputs + 1:numInputs + numOutputs,
variables.lambda(1):variables.lambda(end)) = Y';
            % Assign output slacks
            for i=1:numOutputs
                A(numInputs + i,variables.ySlack(i)) = -1;
            end;
        else
            % We have integer constraints
            % Assign X:
                  A(1:numInputs,variables.lambda(1) : variables.lambda(end)) = -
        X';
            % Assign the intermediate integer variables
            for i=1:numInputs
                A(i,variables.xInteger(i)) = 1;
            end;
            % Assign input slacks (these are separate
            % constraints from the ones above)
            for i=1:numInputs
                A(numInputs + numOutputs + i, variables.xSlack(i)) = -1;
                % Also assign the intermediate integer
                % variables
                A(numInputs+numOutputs+i, variables.xInteger(i)) = -1;
            end;
            % Assign Y:
            A(numInputs + 1:numInputs + numOutputs,
variables.lambda(1):variables.lambda(end)) = Y';
            % Assign the intermediate integer variables
            for i=1:numOutputs
                A(numInputs+i, variables.yInteger(i)) = -1;
            end;
            % Assign output slacks (these are separate
            % constraints from the ones above)
```

```
            for i=1:numOutputs
                A(numInputs+numOutputs+numInputs + i, variables.ySlack(i)) = -1;
                % Also assign the intermediate integer
                % variables
                A(numInputs + numOutputs + numInputs + i,
variables.yInteger(i)) = 1;
            end;
        end;
    else
        if sum(integerParams(:)) == 0
            % Output oriented model
            % Assign X:
                 A(1:numInputs, variables.lambda(1) : variables.lambda(end)) =
        X';
            % Assign input slacks
            for i=1:numInputs
                A(i,variables.xSlack(i)) = 1;
            end;
            % Assign Y:
            A(numInputs + 1:numInputs + numOutputs, variables.lambda(1) :
variables.lambda(end)) = -Y';
            % Assign output slacks
            for i=1:numOutputs
                A(numInputs+ i, variables.ySlack(i)) = 1;
            end;
        else
            % We have integer variables
            % Output oriented model
            % Assign X:
                       A(1:numInputs, variables.lambda(1) :
             variables.lambda(end)) = X';
            % Assign the intermediate integer variables
            for i=1:numInputs
```

```
                A(i,variables.xInteger(i)) = -1;
            end;
            % Assign input slacks (these are separate
            % constraints from the ones above)
            for i=1:numInputs
                    A(numInputs + numOutputs + i, variables.xSlack(i)) = 1;
                    A(numInputs + numOutputs + i, variables.xInteger(i)) = 1;
            end;
            % Assign Y:
                 A(numInputs + 1:numInputs + numOutputs, variables.lambda(1) :
        variables.lambda(end)) = -Y';
            % Assign the intermediate integer variables
            for i=1:numOutputs
                A(numInputs +i, variables.yInteger(i)) = 1;
            end;
            % Assign output slacks (these are separate
            % constraints from the ones above)
            for i=1:numOutputs
                A(numInputs+numOutputs + numInputs + i,variables.ySlack(i)) = 1;
                % Also assign the intermediate integer
                % variables
                A(numInputs+numOutputs+numInputs+i,
                  variables.yInteger(i)) = -1;
            end;
        end;
    end;
  end;
end


function F = setF(numDMUs, numInputs, numOutputs, DEAmodel, orientation, weightlbs,
      DMU, X, Y, variables, numVars)
    % Set the objective function:
```

```
    F = zeros(numVars,1)';
    if strcmp(DEAmodel,'BCC') || strcmp(DEAmodel,'CCR')
        %this is theta
        F(1,1) = variables.theta;
        if strcmp(orientation, 'input')
            F(variables.xSlack(1):variables.ySlack(end)) = -weightlbs;
        else
            F(variables.xSlack(1):variables.ySlack(end)) = weightlbs;
        end;
    elseif strcmp(DEAmodel,'SBM')
        %this is t
        F(1) = variables.theta;
        for i=1:numInputs
            F(variables.xSlack(i)) = -1/(numInputs * X(DMU,i));
        end;
    elseif strcmp(DEAmodel,'ADDNEG')
        for i = 1 :  numInputs
            coeff = (numInputs+numOutputs)*(X(DMU,i) -
min(X(:,i)));
                if coeff ~= 0
                    F(variables.xSlack(i)) = 1/coeff;
                else
                    F(variables.xSlack(i)) = 0;
                end;
        end;
        for i = 1 : numOutputs
            coeff = (numInputs+numOutputs)*(max(Y(:,i)) - Y(DMU,i));
            if coeff ~= 0
                F(variables.ySlack(i)) = 1/coeff;
            else
                F(variables.ySlack(i)) = 0;
            end;
```

```
        end;
    end;
end

function Fphase2 = setFphase2(numDMUs, numInputs, numOutputs,
     orientation, variables, numVars)
    % Set the phase2 objective function
    Fphase2 = zeros(numVars,1)';
    if strcmp(orientation,'input')
        Fphase2(variables.xSlack(1):variables.xSlack(end)) =-1;
        Fphase2(variables.ySlack(1):variables.ySlack(end)) =-1;
    else
        Fphase2(variables.xSlack(1):variables.xSlack(end)) = 1;
        Fphase2(variables.ySlack(1):variables.ySlack(end)) = 1;
    end;
end
```

B.2 Code for Batch Execution of DEA Run

```
function batchDEA(outfileName, dataFiles, weightsLbs, numInputs, numOutputs,
     integerParams, DEAmodel, orientation)
%This function calls DEA repeatedly for a large set of files.
% Inputs:
%    - outfileName: The file to which results should be printed
%    - weightsLbs: The lower bound on weights for each run, indexed by i (run) and j
(variable)
%    - dataFiles: The list of datafiles to be used, indexed by i (run)
%    - The number of inputs
%    - The number of outputs
%    - An integer matrix (numInputs+numOutputs,1) with 1s marking those
inputs/outputs that are integer, and 0s otherwise.
%    - The DEA model.  Current choices are (case sensitive): CCR, BCC, SBM, ADDNEG
%    - The orientation of the DEA model ('input' or 'output')
%    - The lower bounds on the variable weights (usually set to 0 or epsilon)
    %set up outfile
    outfile = fopen(outfileName,'w');
    for run=1:length(dataFiles)
            % set print mode to 2 in the first loop and then 3 subsequently
            DEA(dataFiles(i), outfile, DEAmodel, orientation, weightLbs(i,:),
     min(i+1,3), outfile);
        end
    end
    fclose(outfile);
end
```

Appendix C: C++ DEA Code

```
  #include <ilcplex/ilocplex.h>
  #include <iostream>
  #include <string>
  #include <fstream>
  #include <sstream>
  #include <vector>
  #include <math.h>
  #include <float.h>
  using namespace std;

  #ifndef EPSILON
  #define EPSILON 0.000001
  #endif

  ILOSTLBEGIN

  void readCSV(istream &input, vector< vector< double> > &output, vector<
string> &DMUnames, vector< string> &varNames, vector< int> &inputVars, vector<
int> &outputVars);

  void printData(vector< vector< double> > &output, vector< string> &DMUnames,
vector< string> &varNames, vector< int> &inputVars, vector< int> &outputVars);

  int populateCCRProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c);
```

```
  int populateBCCProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c);


  int populateSBMProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c);


  void resultsPrinter(vector< double> &effScores, vector< string>
&DMUnameListForResults, vector< vector< double> > &paramVals, int numParams,
int numInputs, vector< string> &DMUnames);


  void resultsPrinterSBM(vector< double> &effScores, vector< string>
&DMUnameListForResults, vector< vector< double> > &paramVals, int numParams,
int numInputs, vector< string> &DMUnames);


  int main (int argc, char* argv[]) {
    if (argc != 4 ) {
      cout << "Incorrect usage.  Please type ./DEA <input filename> <target
DMU> <DEA model>\n\n<target DMU> can be a numeric value or 'all' in order to
calculate efficiencies for all DMUs.\n<DEA model can be: 'CCR','BCC', or
'SBM'.\n  Only input-oriented models implemented.";
      return 0;
    }
    else {
      fstream file(argv[1], ios::in);
      if(!file.is_open())
       {
         cout << "File not found!\n";
         return 0;
       }
      // typedef to save typing for the following object
      typedef vector< vector< double> > csvVector;
      csvVector csvData;

      //list of DMU names
      vector< string> DMUnames;

      //list of input and output variable names
      vector< string> varNames;
```

```
	//list of input variable locations
	vector< int> inputVars;

	//list of output variable locations
	vector< int> outputVars;

	//total number of variables in problem
	int totalNumVars;

	//list of results
	vector< double> effScores;
	vector< string> DMUnameListForResults;
	vector< vector< double> > paramVals;
	//can add more here as necessary

	readCSV(file, csvData, DMUnames, varNames, inputVars, outputVars);

	// this line can be deleted but should be moved to error message for when
the read-in doesn't work
	// printData(csvData, DMUnames, varNames, inputVars, outputVars);

	IloEnv env;

	int startingDMU, endingDMU;
	if (strcmp(argv[2], "all") == 0) {
	 startingDMU = 0;
	 endingDMU = DMUnames.size() - 1;
	}
	else {
	 startingDMU = atoi(argv[2]);
	 endingDMU = startingDMU;
	}
	for(int i=startingDMU; i <= endingDMU; i++) {
	 try {
	   IloModel model(env);
	   IloNumVarArray var(env);
```

```
        IloRangeArray con(env);

        // populate problem
        if (strcmp(argv[3], "CCR") == 0) {
          totalNumVars = populateCCRProblem(csvData, inputVars, outputVars, i,
model, var, con);
        }
        else if (strcmp(argv[3], "BCC") == 0) {
          totalNumVars = populateBCCProblem(csvData, inputVars, outputVars, i,
model, var, con);
        }
        else if (strcmp(argv[3], "SBM") == 0) {
          totalNumVars = populateSBMProblem(csvData, inputVars, outputVars, i,
model, var, con);
        }
        else {
          cout << "Incorrect DEA model specified.  Options are: CCR, BCC, or
SBM\n";
          return 0;
        }

        IloCplex cplex(model);

        // turn off cplex output to screen
        cplex.setOut(env.getNullStream());

        // Optimize the problem and obtain solution.
        if ( !cplex.solve() ) {
          env.error() << "Failed to optimize LP" << endl;
          throw(-1);
        }

        // writing results
        IloNumArray vals(env);
        effScores.push_back(cplex.getObjValue());
        DMUnameListForResults.push_back(DMUnames[i]);
        cplex.getValues(vals, var);
        vector< double> tempVals;
```

```
          for (int j = 0; j < totalNumVars; j++) {
            tempVals.push_back(vals[j]);
          }
          paramVals.push_back(tempVals);

          cplex.exportModel("lpex1.lp");
        }
        catch (IloException& e) {
          cerr << "Concert exception caught: " << e << endl;
        }
        catch (...) {
          cerr << "Unknown exception caught" << endl;
        }
      }
      // print results
      if (strcmp(argv[3], "SBM") == 0) {
        resultsPrinterSBM(effScores, DMUnameListForResults, paramVals,
totalNumVars, inputVars.size(), DMUnames);
      }
      else {
        resultsPrinter(effScores, DMUnameListForResults, paramVals, totalNumVars,
inputVars.size(), DMUnames);
      }
      env.end();

    }
    return 0;
  }

  void resultsPrinter(vector< double> &effScores, vector< string>
&DMUnameListForResults, vector< vector< double> > &paramVals, int numParams,
int numInputs, vector< string> &DMUnames) {

    cout << "\n-------------\n-- Results --\n-------------\n\n";
    cout << "DMU name\tScore\t";
    // print header for lambdas
    for (int i = 0; i < DMUnames.size(); i++) {
      cout << "La" << DMUnames[i].substr(0,4) << "\t";
```

```
    }
    for (int i = 0; i < numInputs; i++) {
      cout << "s" << i+1 << "-\t";
    }
    for (int i = 0; i < numParams - numInputs - DMUnames.size() - 1; i++) {
      cout << "s" << i+1 << "+\t";
    }
    cout << "\n";
    for (int i = 0; i < numParams + 2; i++) {
      cout << "--------";
    }
    cout << "\n";
    for(int i = 0; i < effScores.size(); i++) {
      cout << DMUnameListForResults[i] << "\t\t" <<
round(effScores[i]*100000)/100000 << "\t";
      for(int j = 1; j < numParams; j++) {
        cout << round(paramVals[i][j]*100000)/100000 << "\t";
      }
      cout << "\n";
    }
  }

  void resultsPrinterSBM(vector< double> &effScores, vector< string>
&DMUnameListForResults, vector< vector< double> > &paramVals, int numParams,
int numInputs, vector< string> &DMUnames) {
    double t;

    cout << "\n-------------\n-- Results --\n-------------\n\n";
    cout << "DMU name\tScore\tt\t";
    // print header for lambdas
    for (int i = 0; i < DMUnames.size(); i++) {
      cout << "La" << DMUnames[i].substr(0,4) << "\t";
    }
    for (int i = 0; i < numInputs; i++) {
      cout << "s" << i+1 << "-\t";
    }
    for (int i = 0; i < numParams - numInputs - DMUnames.size() - 1; i++) {
```

```
      cout << "s" << i+1 << "+\t";
    }
    cout << "\n";
    for (int i = 0; i < numParams + 3; i++) {
      cout << "--------";
    }
    cout << "\n";
    for(int i = 0; i < effScores.size(); i++) {
      t = paramVals[i][0];
      cout << DMUnameListForResults[i] << "\t\t" <<
round(effScores[i]*100000)/100000 << "\t";
      cout << round(paramVals[i][0]*100000)/100000 << "\t";
      for(int j = 1; j < numParams; j++) {
       cout << round(paramVals[i][j]/t*100000)/100000 << "\t";
      }
      cout << "\n";
    }
  }

  void readCSV(istream &input, vector< vector< double> > &output, vector<
string> &DMUnames, vector< string> &varNames, vector< int> &inputVars, vector<
int> &outputVars) {
    string csvLine;
    // read in input and output variable locations
    getline(input, csvLine);
    istringstream csvStream(csvLine);
    string csvElement;

    //clean out the first empty cell
    getline(csvStream, csvElement, ',');
    // read every element from the line that is separated by commas
    // and put it into the vector or strings
    int colCounter = 1;
    while( getline(csvStream, csvElement, ',') ){
      if(csvElement == string("i")) {
       inputVars.push_back(colCounter);
      }
```

```
    else {
     outputVars.push_back(colCounter);
    }
    colCounter++;
  }

  // read in variable names
  getline(input, csvLine);
  istringstream csvStream2(csvLine);

  //clean out the first empty cell
  getline(csvStream2, csvElement, ',');
  // read every element from the line that is separated by commas
  // and put it into the vector or strings
  while( getline(csvStream2, csvElement, ',') ){
    varNames.push_back(csvElement);
  }

  // read every remaining line from the stream
  while( getline(input, csvLine) )
    {
     istringstream csvStream(csvLine);
     vector<double> csvColumn;
     string csvElement;
     // read every element from the line that is separated by commas
     // and put it into the vector or strings
     getline(csvStream, csvElement, ',');
     DMUnames.push_back(csvElement);
     while( getline(csvStream, csvElement, ',') ){
       csvColumn.push_back(atof(csvElement.c_str()));
     }
     output.push_back(csvColumn);
    }
}
```

```
  int populateCCRProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c) {
    IloEnv env = model.getEnv();
    IloObjective obj = IloMinimize(env);

    // number of variables for CCR dual is 1 (theta) + numDMUs (rows in output)
+ numInputs (length of inputVars) + numOutputs (length of outputVars)
    int numVars = 1 + output.size() + inputVars.size() + outputVars.size();
    for(int i = 0; i < numVars; i++) {
      x.add(IloNumVar(env));
    }

    for(int i = 0; i< inputVars.size(); i++) {
      c.add(IloRange(env, 0.0, 0.0));
    }
    for(int i = 0; i< outputVars.size(); i++) {
      double rhsValue = output[goalDMU][outputVars[i] - 1];
      c.add(IloRange(env, rhsValue, rhsValue));
    }

    obj.setLinearCoef(x[0], 1.0);

    //count which constraint row we're on
    int constraintCounter = 0;

    //write input constraints
    int varCounter;
    int varOffset = 0;
    for(vector< int>::iterator j = inputVars.begin(); j != inputVars.end();
++j) {
      varCounter = 0;
      c[constraintCounter].setLinearCoef(x[0], output[goalDMU][*j - 1]);
      varCounter++;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
        c[constraintCounter].setLinearCoef(x[varCounter], -i[0][*j - 1]);
        varCounter++;
```

```
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], -1);
      constraintCounter++;
      varOffset++;
    }

    //write output constraints
    for(vector< int>::iterator j = outputVars.begin(); j != outputVars.end();
++j) {
      varCounter = 1;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
        c[constraintCounter].setLinearCoef(x[varCounter], i[0][*j - 1]);
        varCounter++;
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], -1.0);
      constraintCounter++;
      varOffset++;
    }

    model.add(obj);
    model.add(c);
    return numVars;
  }

  int populateBCCProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c) {
    IloEnv env = model.getEnv();
    IloObjective obj = IloMinimize(env);

    // number of variables for BCC dual is 1 (theta) + numDMUs (rows in output)
+ numInputs (length of inputVars) + numOutputs (length of outputVars)
    int numVars = 1 + output.size() + inputVars.size() + outputVars.size();
    for(int i = 0; i < numVars; i++) {
      x.add(IloNumVar(env));
    }
```

```
    for(int i = 0; i< inputVars.size(); i++) {
      c.add(IloRange(env, 0.0, 0.0));
    }
    for(int i = 0; i< outputVars.size(); i++) {
      double rhsValue = output[goalDMU][outputVars[i] - 1];
      c.add(IloRange(env, rhsValue, rhsValue));
    }

    //add room for the unity constraint
    c.add(IloRange(env, 1.0, 1.0));

    obj.setLinearCoef(x[0], 1.0);



    //count which constraint row we're on
    int constraintCounter = 0;

    //write input constraints
    int varCounter;
    int varOffset = 0;
    for(vector< int>::iterator j = inputVars.begin(); j != inputVars.end();
++j) {
      varCounter = 0;
      c[constraintCounter].setLinearCoef(x[0], output[goalDMU][*j - 1]);
      varCounter++;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
       c[constraintCounter].setLinearCoef(x[varCounter], -i[0][*j - 1]);
       varCounter++;
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], -1);
      constraintCounter++;
      varOffset++;
    }

    //write output constraints
```

```
    for(vector< int>::iterator j = outputVars.begin(); j != outputVars.end();
++j) {
      varCounter = 1;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
       c[constraintCounter].setLinearCoef(x[varCounter], i[0][*j - 1]);
       varCounter++;
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], -1.0);
      constraintCounter++;
      varOffset++;
    }

    //write unity constraint in BCC model
    for(varCounter = 1; varCounter <= output.size(); varCounter++) {
      c[constraintCounter].setLinearCoef(x[varCounter], 1.0);
    }

    model.add(obj);
    model.add(c);
    return numVars;
  }

  int populateSBMProblem(vector< vector< double> > &output, vector< int>
&inputVars, vector< int> &outputVars, int goalDMU, IloModel model,
IloNumVarArray x, IloRangeArray c) {
    IloEnv env = model.getEnv();
    IloObjective obj = IloMinimize(env);

    // number of variables for dual is 1 (theta) + numDMUs (rows in output) +
numInputs (length of inputVars) + numOutputs (length of outputVars)
    int numVars = 1 + output.size() + inputVars.size() + outputVars.size();
    for(int i = 0; i < numVars; i++) {
      x.add(IloNumVar(env));
    }

    c.add(IloRange(env, 1.0, 1.0));
```

```
    for(int i = 0; i< inputVars.size(); i++) {
      c.add(IloRange(env, 0.0, 0.0));
    }
    for(int i = 0; i< outputVars.size(); i++) {
      c.add(IloRange(env, 0.0, 0.0));
    }

    // SBMt objective
    obj.setLinearCoef(x[0], 1.0);
    int objVarCounter = output.size();
    for(int i = 0; i< inputVars.size(); i++) {
      //      cout << "inputVars.size(): " << inputVars.size() << "
output[goalDMU][inputVars[i] - 1]: " << output[goalDMU][inputVars[i] - 1] << "
Res: " << (-1.0/inputVars.size()) * (1.0/output[goalDMU][inputVars[i] - 1]) <<
"\n";
      obj.setLinearCoef(x[objVarCounter + 1], (-1.0/inputVars.size()) *
(1.0/output[goalDMU][inputVars[i] - 1]));
      objVarCounter++;
    }

    //count which constraint row we're on
    int constraintCounter = 0;

    // write initial t constraint
    int initialVarCounter = output.size() + inputVars.size();
    c[constraintCounter].setLinearCoef(x[0], 1.0);
    for(int i = 0; i< outputVars.size(); i++) {
      c[constraintCounter].setLinearCoef(x[initialVarCounter + 1],
1.0/outputVars.size()*1.0/output[goalDMU][outputVars[i] - 1]);
      initialVarCounter++;
    }
    constraintCounter++;

    //write input constraints
    int varCounter;
    int varOffset = 0;
    for(vector< int>::iterator j = inputVars.begin(); j != inputVars.end();
++j) {
```

```
      varCounter = 0;
      c[constraintCounter].setLinearCoef(x[0], output[goalDMU][*j - 1]);
      varCounter++;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
       c[constraintCounter].setLinearCoef(x[varCounter], -i[0][*j - 1]);
       varCounter++;
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], -1);
      constraintCounter++;
      varOffset++;
    }

    //write output constraints
    for(vector< int>::iterator j = outputVars.begin(); j != outputVars.end();
++j) {
      c[constraintCounter].setLinearCoef(x[0], output[goalDMU][*j - 1]);
      varCounter = 1;
      for(vector< vector< double> >::iterator i = output.begin(); i !=
output.end(); ++i) {
       c[constraintCounter].setLinearCoef(x[varCounter], -i[0][*j - 1]);
       varCounter++;
      }
      c[constraintCounter].setLinearCoef(x[varCounter + varOffset], 1.0);
      constraintCounter++;
      varOffset++;
    }

    // workaround to specify variable t > 0
    c.add(x[0] >= EPSILON);

    model.add(obj);
    model.add(c);
    return numVars;
  }
```

Appendix D: Stakeholder Knowledge Elicitation

D.1 Subject-Matter Experts

Table D.1 presents the participating subject-matter experts. The knowledge elicitation sessions took place in 2009 and were conducted by David Schaar in-person and via phone.

Table D.1 – Participants in knowledge elicitation sessions on the definition of airport stakeholders, their goals, and key performance indicators for airports.

Person	Title	Organization	Date
Jim Bennett	President and Chief Executive Officer	Metropolitan Washington Airports Authority	16 Oct, 2009
Frank Berardino	President	GRA, Inc.	10 Nov, 2009
Les Berry	Director of Strategic Planning	DEN Airport	18 Nov, 2009

Person	Title	Organization	Date
Tom Bock	General manager for airspace technology and airspace enhancements	Port Authority of New York New Jersey	22 Apr, 2009
Chellie Cameron	Manager of Financial Strategy and Research	Metropolitan Washington Airports Authority	27 Oct, 2009
Michael Cintron	Head of advocacy and traveler consumer affairs	International Air Passengers Association	20 Oct, 2009
Ken Cushine	Vice President	Frasca and Associates	9 Nov, 2009
Dallas Dawson	Performance Management Analyst	TPA Airport	19 Oct, 2009
Lorie Dewitt	Hub supervisor	United Airlines	1 Jul, 2009
Matt Erskine	Executive Director	Greater Washington Board of Trade	4 Nov, 2009
Kurt Forsgren	Vice President of credit analysis for public transportation infrastructure	Standard and Poor's	28 Oct, 2009
Stephen Freibrun	Head of advisory practice on non-aeronautical revenue for airports	SH&E / ICF International	19 Oct, 2009
Jeff Gilley	Airports Operations Division	National Business Aviation Association	3 Nov, 2009

Person	Title	Organization	Date
Dilwyn Gruffydd	Project Manager	Landrum Brown	5 Nov, 2009
Liying Gu	Senior Director of Economic Affairs and Research	Airports Council International - North America	21 Oct, 2009
Susan Kopinski	Chief Financial Officer and deputy airports director for finance and administration	STL Airport	10 Nov, 2009
Kurt Krummenacker	Vice President, Global Infrastructure and Project Finance	Moody's	16 Oct, 2009
Seth Lehman and Emma Walker	Senior Director and Associate Director, respectively	Fitch Ratings	10 Nov, 2009
Peter Mackenzie Williams	Associate Director	Jacobs Consultancy	5 Nov, 2009
Laura McKee	Managing Director - Airport Affairs	Air Transport Association	11 Nov, 2009
Paul McKnight	Associate Director	Jacobs Consultancy	3 Nov, 2009
Pat Oldfield	Manager of operations analysis group	United Airlines	20 Oct, 2009
Bob O'Roark	Design task manager	Metropolitan Washington Airports Authority	12 Jun, 2009

Person	Title	Organization	Date
Jake Plante	Noise and air quality resource expert	FAA Airports Office	3 Nov, 2009
Chris Poinsatte	Executive Vice President and Chief Financial Officer	DFW Airport	16 Nov, 2009
Theresia Schatz	Senior Program Officer	Airport Cooperative Research Program	22 Apr, 2009
Peter Stettler	Director	Ricondo and Associates	11 Nov, 2009
Terry Thompson	Vice President and Chief Environmental Scientist	Metron Aviation	29 Oct, 2009
Jim Walsh	Vice President	Landrum Brown	5 Nov, 2009
Jim Wilding	Former President and Chief Executive Officer	Metropolitan Washington Airports Authority	19 Apr, 2009
Gregg Wollard	Aviation Planner	Metropolitan Washington Airports Authority	20 Mar, 2009
Alex Zaslov	Senior Aviation Consultant	HNTB Inc.	25 Jun, 2009

D.2 Knowledge Elicitation Form

Each knowledge elicitation session followed the same structure, using the questions listed below:

1. What is your position?

2. What do you consider the goals and objectives for the airport to be?

3. Who do you consider to be the main constituents of the airport?

4. Which performance metrics do you consider to be the most important for the airport?

5. What do these metrics tell you and which decisions do they support?

6. Are there any metrics that you don't currently track that you think would help gauge the airport's performance?

7. Who else would you recommend I speak with about this topic?

Appendix E: Detailed Benchmark Data and Results

This appendix provides details of the inputs and outputs used in each benchmark, and details of each benchmark's results. The data sources, as described in section 3.4, include:

- **Data on airline service**: Data on the traffic between airport pairs was obtained from the T100 database which is compiled from data collected by Office of Airline Information (OAI) at the Bureau of Transportation Statistics (BTS) (Bureau of Transportation Statistics 2010b). It included variables such as the frequency of service, the available seat capacity, and the number of passengers carried.
- **Airfare data**: Data on airfares was obtained from the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c).

- **Airport financial data**: Data on airport revenues and costs was obtained from the FAA's Compliance Activity Tracking System provides (Federal Aviation Administration 2010a).
- **Aircraft movement volume data**: Data on aircraft movements was obtained from the FAA's Air Traffic Activity System (Federal Aviation Administration 2010)
- **Data on on-time performance**: On-time data was compiled from data collected by the OAI at the BTS (Bureau of Transportation Statistics 2010b).
- **GDP data**: Data on GDP by metropolitan area was obtained from the U.S. government's Bureau of Economic Analysis (BEA) (Bureau of Economic Analysis, U.S. Department of Commerce 2010).
- **Population data**: Data on the population by metropolitan area was obtained from the U.S. Census Bureau (U.S. Census Bureau 2010b).
- **Runway capacity**: This data was derived from the analysis described in (Kumar & Sherry 2009). This analysis in turn was conducted using the Aviation System Performance Metrics (ASPM) database (Federal Aviation Administration 2010d) along with the T100 database

described above and the Airline Origin and Destination Survey (DB1B) database (Bureau of Transportation Statistics 2010c).

- **Runways**: This data was compiled from the FAA's National Plan of Integrated Airport Systems (FAA 2008).

E.1 Case Study 1: Benchmarking the Level of Domestic Air Service to U.S. Metropolitan Areas

E.1.1 Benchmark Data

Table E.1 - Case study 1 benchmark data

Metropolitan area	Population	GDP (million current US$)	Number of non-hub nonstop destinations	Number of departures to top 5 hubs
2005				
Atlanta	4,945,773	243,020	158	110
Boston	6,463,090	336,653	72	79
Charlotte	1,518,488	101,877	94	59
Chicago	9,390,691	459,013	135	159
Cincinnati	2,944,294	122,942	127	82
Cleveland	2,116,304	98,109	73	50
Dallas	5,817,696	315,710	138	113
Denver	2,358,271	131,072	120	93
Detroit	4,496,480	199,441	124	66
Honolulu	899,673	41,295	27	9
Houston	5,302,908	312,314	119	97
Las Vegas	1,702,957	83,153	99	77
Los Angeles	16,686,936	738,967	76	139
Memphis	1,252,785	57,419	83	49
Miami	5,375,410	233,824	64	98
Minneapolis	3,131,632	172,356	140	99
New York	18,812,037	1,055,344	103	175
Orlando	1,935,502	90,129	82	63
Philadelphia	5,786,636	295,454	88	76

Metropolitan area	Population	GDP (million current US$)	Number of non-hub nonstop destinations	Number of departures to top 5 hubs
Phoenix	3,873,404	163,452	89	75
Pittsburgh	2,372,356	100,018	77	44
Portland	2,087,066	93,734	42	32
Salt Lake City	1,045,905	51,386	89	50
San Diego	2,931,689	147,733	34	42
San Francisco	5,900,486	396,088	62	97
Seattle	3,197,370	183,671	66	51
St. Louis	2,773,156	115,125	79	71
Tampa	2,637,036	100,907	68	42
Washington-Baltimore	7,868,885	463,723	102	179
2006				
Atlanta	5,113,924	255,382	173	105
Boston	6,470,867	355,331	71	76
Charlotte	1,580,070	113,498	94	58
Chicago	9,439,805	488,255	136	164
Cincinnati	2,960,941	126,019	117	75
Cleveland	2,103,850	100,073	73	49
Dallas	5,995,596	340,639	133	113
Denver	2,403,113	138,450	125	93
Detroit	4,486,620	198,513	121	72
Honolulu	904,134	44,263	28	8
Houston	5,485,545	346,338	117	95
Las Vegas	1,770,676	89,881	115	81
Los Angeles	16,779,490	787,499	81	134
Memphis	1,270,263	60,381	81	49
Miami	5,402,334	251,501	60	96
Minneapolis	3,164,180	178,479	130	95
New York	18,848,240	1,134,178	101	182
Orlando	1,993,945	97,837	82	62
Philadelphia	5,805,349	309,977	87	70
Phoenix	4,035,176	179,788	90	83
Pittsburgh	2,360,750	105,459	69	39
Portland	2,126,475	104,426	43	33
Salt Lake City	1,074,254	56,260	99	55
San Diego	2,937,023	155,458	33	42
San Francisco	5,939,869	422,805	64	95
Seattle	3,253,977	195,167	68	55
St. Louis	2,790,760	117,833	75	67
Tampa	2,687,091	108,159	63	40
Washington-Baltimore	7,917,819	486,633	99	164
2007				
Atlanta	5,261,296	267,295	175	102
Boston	6,491,746	372,977	76	74
Charlotte	1,646,431	116,501	96	58

Metropolitan area	Population	GDP (million current US$)	Number of non-hub nonstop destinations	Number of departures to top 5 hubs
Chicago	9,496,853	510,666	147	150
Cincinnati	2,982,652	129,878	109	76
Cleveland	2,094,885	102,956	73	46
Dallas	6,153,474	362,075	131	107
Denver	2,453,393	143,914	134	92
Detroit	4,457,523	200,742	122	70
Honolulu	900,525	46,358	28	8
Houston	5,597,958	375,451	116	94
Las Vegas	1,827,655	95,737	124	78
Los Angeles	16,851,185	811,689	87	139
Memphis	1,279,120	62,953	78	48
Miami	5,392,118	260,043	62	98
Minneapolis	3,197,620	186,738	129	89
New York	18,922,571	1,209,997	107	184
Orlando	2,028,669	102,118	83	66
Philadelphia	5,823,285	322,325	84	67
Phoenix	4,165,921	186,577	91	80
Pittsburgh	2,354,159	110,489	64	38
Portland	2,166,491	109,637	46	30
Salt Lake City	1,095,362	60,594	101	54
San Diego	2,959,734	162,118	44	42
San Francisco	6,002,480	448,161	67	95
Seattle	3,298,225	210,364	70	55
St. Louis	2,805,465	122,096	69	69
Tampa	2,715,273	110,743	61	42
Washington-Baltimore	7,966,100	508,455	100	151
2008				
Atlanta	5,376,285	269,799	162	101
Boston	6,521,511	385,524	76	68
Charlotte	1,701,799	118,350	104	59
Chicago	9,569,624	520,672	146	143
Cincinnati	2,991,681	132,528	105	73
Cleveland	2,088,291	104,425	84	42
Dallas	6,300,006	379,863	134	107
Denver	2,506,626	150,810	142	88
Detroit	4,425,110	200,856	124	69
Honolulu	905,034	48,095	24	7
Houston	5,728,143	403,202	115	90
Las Vegas	1,865,746	97,053	118	71
Los Angeles	16,988,679	830,964	87	133
Memphis	1,285,732	63,826	81	47
Miami	5,414,772	261,263	77	99
Minneapolis	3,229,878	193,947	134	89
New York	19,006,798	1,264,896	106	180

Metropolitan area	Population	GDP (million current US$)	Number of non-hub nonstop destinations	Number of departures to top 5 hubs
Orlando	2,054,574	103,985	82	63
Philadelphia	5,838,471	331,897	88	69
Phoenix	4,281,899	187,431	87	74
Pittsburgh	2,351,192	114,707	42	34
Portland	2,207,462	112,420	47	30
Salt Lake City	1,115,692	62,525	95	56
San Diego	3,001,072	169,325	47	44
San Francisco	6,093,729	457,512	68	87
Seattle	3,344,813	218,771	71	56
St. Louis	2,816,710	128,467	67	68
Tampa	2,733,761	110,510	64	41
Washington-Baltimore	8,025,247	528,759	104	151

E.1.2 Results

Table E.2 - Case study 1 benchmark results

Metropolitan area	Results 2005	2006	2007	2008
Atlanta	1.0000	1.0000	1.0000	1.0000
Boston	1.7847	1.7827	1.7405	1.8013
Charlotte	1.0681	1.1768	1.2324	1.1039
Chicago	1.0000	1.0000	1.0000	1.0000
Cincinnati	1.0000	1.0912	1.1560	1.1282
Cleveland	1.5082	1.6243	1.7280	1.4440
Dallas	1.0870	1.1024	1.1209	1.0872
Denver	1.0000	1.0000	1.0000	1.0000
Detroit	1.1847	1.2650	1.2723	1.2130
Honolulu	1.0000	1.0000	1.0000	1.0000
Houston	1.2368	1.2540	1.2436	1.2564
Las Vegas	1.0000	1.0000	1.0000	1.0000
Los Angeles	1.2857	1.2829	1.1913	1.2186
Memphis	1.1005	1.1746	1.1568	1.1835
Miami	1.2184	1.2086	1.1247	1.0763
Minneapolis	1.0000	1.0333	1.0814	1.0572
New York	1.0155	1.0000	1.0000	1.0000
Orlando	1.2540	1.3359	1.2166	1.1520
Philadelphia	1.6548	1.7260	1.7502	1.6447
Phoenix	1.3539	1.2272	1.2408	1.2619
Pittsburgh	1.4913	1.7458	2.0060	2.2129
Portland	2.5082	2.5867	2.6919	2.5386
Salt Lake City	1.0000	1.0000	1.0000	1.0000
San Diego	2.3335	2.3071	2.2573	2.0747
San Francisco	1.5326	1.4515	1.3636	1.4704
Seattle	2.0001	1.8626	1.8468	1.7411
St. Louis	1.2462	1.3084	1.2429	1.1856
Tampa	1.6956	1.9271	1.9392	1.8327
Washington-Baltimore	1.0000	1.0000	1.0000	1.0000

E.2 Case Study 2: Benchmark of the Level of Capacity Utilization at U.S. Airports

E.2.1 Benchmark Data

Table E.3 - Case study 2 benchmark data

Airport	Number of non-hub nonstop destinations	Number of departures to top 5 hubs	Enplaned domestic passengers annually	Capacity for domestic passenger flights (number of flights per 15 minutes)
2005				
ATL	158	110	38,977,547	47.82
BOS	71	55	11,275,847	21.97
BWI	59	47	9,950,019	18.01
CLE	73	50	5,448,414	21.79
CLT	94	59	13,343,785	28.36
CVG	126	56	10,828,255	44.21
DCA	69	67	8,560,309	16.46
DEN	120	93	20,315,605	55.19
DFW	138	113	25,853,877	45.16
DTW	124	66	15,877,941	34.18
EWR	84	73	11,870,774	16.07
FLL	57	38	9,853,012	15.88
HNL	27	9	7,791,782	22.24
IAD	84	65	10,753,502	29.24
IAH	118	77	15,798,725	33.99
JFK	56	17	10,998,696	12.54
LAS	99	77	21,025,158	25.60
LAX	75	81	21,714,443	28.69
LGA	65	81	12,268,482	16.89
MCO	82	63	15,672,082	36.97
MDW	59	44	8,727,436	16.81
MEM	83	49	5,537,595	20.87
MIA	43	44	7,830,714	15.19
MSP	140	99	16,843,327	32.35
ORD	130	115	31,596,042	40.53
PDX	42	32	6,669,393	23.71
PHL	88	76	13,876,702	21.17
PHX	89	75	20,358,777	31.64
PIT	77	44	5,193,210	36.16
SAN	34	42	8,629,842	11.34

Airport	Number of non-hub nonstop destinations	Number of departures to top 5 hubs	Enplaned domestic passengers annually	Capacity for domestic passenger flights (number of flights per 15 minutes)
SEA	66	51	13,453,176	18.99
SFO	57	56	12,472,436	18.38
SLC	89	50	10,669,931	32.08
STL	79	71	7,104,506	29.11
TPA	68	42	9,280,042	26.20
2006				
ATL	173	105	37,435,075	47.36
BOS	71	55	11,677,157	22.17
BWI	55	46	10,549,699	17.98
CLE	73	49	5,369,092	21.78
CLT	94	58	13,991,100	28.42
CVG	116	48	7,625,303	45.14
DCA	67	67	8,901,154	16.47
DEN	125	93	22,132,888	55.16
DFW	133	113	26,396,788	45.05
DTW	121	72	15,846,969	34.14
EWR	80	70	12,809,318	15.80
FLL	50	38	9,314,653	15.61
HNL	28	8	7,852,255	22.68
IAD	80	51	8,600,046	27.83
IAH	116	74	16,952,946	34.06
JFK	59	18	11,481,445	12.67
LAS	115	81	21,611,422	25.62
LAX	80	82	21,778,244	28.87
LGA	69	91	12,279,330	16.96
MCO	82	62	15,971,444	37.04
MDW	54	50	9,329,558	16.86
MEM	81	49	5,412,395	20.63
MIA	44	43	8,355,291	15.10
MSP	130	95	16,128,647	32.15
ORD	133	114	31,493,226	40.25
PDX	43	33	6,814,532	23.78
PHL	87	70	13,936,097	21.22
PHX	90	83	20,690,143	31.69
PIT	69	39	4,975,637	35.91
SAN	33	42	8,703,594	11.38
SEA	68	55	13,827,492	19.17
SFO	57	56	12,436,322	18.23
SLC	99	55	10,292,948	31.69
STL	75	67	7,351,497	29.24
TPA	63	40	9,195,480	26.22
2007				
ATL	175	102	38,950,038	47.23
BOS	76	54	11,881,356	22.02

Airport	Number of non-hub nonstop destinations	Number of departures to top 5 hubs	Enplaned domestic passengers annually	Capacity for domestic passenger flights (number of flights per 15 minutes)
BWI	59	43	10,766,947	18.03
CLE	73	46	5,434,231	21.78
CLT	96	58	15,684,161	28.56
CVG	108	47	7,378,115	45.41
DCA	68	63	8,899,333	16.58
DEN	134	92	23,390,913	54.85
DFW	131	107	26,305,140	45.02
DTW	122	70	15,727,865	34.03
EWR	83	68	12,833,251	15.51
FLL	52	39	9,564,488	15.43
HNL	28	8	8,524,306	22.93
IAD	81	46	9,007,457	27.54
IAH	115	73	16,936,107	33.99
JFK	61	28	12,801,910	13.06
LAS	124	78	21,844,934	25.51
LAX	85	84	22,287,234	28.91
LGA	69	87	11,856,324	16.98
MCO	83	66	16,658,633	37.13
MDW	60	41	9,578,510	16.90
MEM	78	48	5,450,488	20.64
MIA	44	43	8,546,808	15.20
MSP	129	89	15,830,779	32.12
ORD	140	110	31,089,529	40.10
PDX	46	30	7,089,451	23.79
PHL	84	67	13,971,199	21.20
PHX	91	80	20,857,326	31.69
PIT	64	38	4,875,841	35.81
SAN	44	42	9,137,301	11.45
SEA	70	55	14,390,682	19.19
SFO	59	57	13,271,672	18.32
SLC	101	54	10,566,545	31.72
STL	69	69	7,426,714	29.14
TPA	61	42	9,341,561	26.21
2008				
ATL	162	101	39,301,439	47.36
BOS	75	53	11,051,321	22.16
BWI	58	42	10,472,809	18.12
CLE	84	42	5,178,839	21.64
CLT	104	59	16,319,115	28.56
CVG	104	46	6,246,714	45.20
DCA	66	65	8,606,561	16.57
DEN	142	88	23,781,956	55.02
DFW	133	105	24,991,709	45.16
DTW	124	69	15,211,968	33.94

Airport	Number of non-hub nonstop destinations	Number of departures to top 5 hubs	Enplaned domestic passengers annually	Capacity for domestic passenger flights (number of flights per 15 minutes)
EWR	87	66	12,178,200	15.30
FLL	66	40	9,499,787	15.47
HNL	24	7	7,274,736	22.57
IAD	85	44	8,508,942	27.07
IAH	114	69	16,062,085	33.75
JFK	60	29	12,704,955	12.83
LAS	118	71	20,465,085	25.41
LAX	83	82	21,068,641	29.00
LGA	70	84	10,992,362	16.87
MCO	82	63	16,190,173	36.85
MDW	51	36	8,643,396	16.94
MEM	81	47	5,238,874	20.92
MIA	46	43	8,420,971	14.96
MSP	134	89	15,205,998	32.24
ORD	144	107	28,435,305	39.90
PDX	47	30	6,846,649	23.76
PHL	88	69	13,917,484	21.23
PHX	87	74	19,467,741	31.59
PIT	42	34	4,290,700	35.54
SAN	47	44	9,097,505	11.49
SEA	71	56	14,608,747	19.28
SFO	62	54	14,121,045	18.63
SLC	95	56	9,895,444	31.91
STL	67	68	7,009,994	29.20
TPA	64	41	8,903,700	26.16

E.2.2 Results

Table E.4 - Case study 2 benchmark results

Airport	Results			
	2005	2006	2007	2008
ATL	1.0000	1.0000	1.0000	1.0000
BOS	1.4067	1.4359	1.4148	1.4087
BWI	1.4159	1.4215	1.4487	1.4532
CLE	1.4204	1.3878	1.4892	1.2671
CLT	1.3050	1.3029	1.3664	1.2025
CVG	1.2207	1.4405	1.5808	1.5193
DCA	1.1303	1.1581	1.1901	1.1859
DEN	1.2040	1.1857	1.1586	1.1408
DFW	1.0065	1.0105	1.0220	1.0232
DTW	1.1463	1.1383	1.1804	1.1061
EWR	1.0000	1.0000	1.0000	1.0000
FLL	1.2970	1.4112	1.4088	1.3104
HNL	2.3989	2.4598	2.3877	2.7784
IAD	1.4717	1.5113	1.5898	1.4341
IAH	1.2026	1.1856	1.2277	1.1970
JFK	1.0000	1.0000	1.0000	1.0000
LAS	1.0000	1.0000	1.0000	1.0000
LAX	1.0645	1.0896	1.0543	1.0742
LGA	1.0000	1.0000	1.0000	1.0000
MCO	1.6580	1.6532	1.6049	1.6322
MDW	1.4429	1.4665	1.4750	1.6494
MEM	1.2111	1.2003	1.3340	1.2870
MIA	1.4594	1.4740	1.4936	1.4489
MSP	1.0000	1.0000	1.0256	1.0000
ORD	1.0000	1.0000	1.0000	1.0000
PDX	2.5056	2.4827	2.5263	2.3937
PHL	1.0817	1.1231	1.1857	1.1237
PHX	1.2417	1.2120	1.2013	1.2514
PIT	1.8758	2.0646	2.3155	2.9777
SAN	1.0000	1.0000	1.0000	1.0000
SEA	1.1458	1.1673	1.1568	1.1094
SFO	1.1624	1.2168	1.1756	1.1166
SLC	1.5629	1.3253	1.3722	1.4024
STL	1.3775	1.4683	1.4351	1.4155
TPA	1.7185	1.8507	1.9532	1.8551

E.3 Case Study 3: Re-design of an Existing Benchmark

E.3.1 Benchmark Data

E.3.1.1 Original Study

Table E.5 - Case study 3 benchmark data for original study

Airport	Run-ways	Gates	Operating expense (US$ million)	Non-operating expense (US$ million)	On-time perfor-mance	Aero-nautical revenues (US$ million)	Non-aero-nautical revenues (US$ million)	Enplane-ments (million pax)	Air carrier operations	Other operations
2005										
ATL	5	184	94.1	78.4	72.6%	97.1	182.9	42.6	692,165	288,221
BOS	6	102	208.8	62.5	75.5%	190.1	177.8	13.2	217,775	203,650
BWI	4	78	123.0	23.4	79.7%	67.4	51.8	10.2	216,520	87,338
CLE	4	68	67.0	32.8	81.5%	69.6	41.5	5.6	80,978	178,693
CLT	3	85	48.7	34.2	80.1%	57.0	45.2	14.3	258,693	264,542
CVG	4	87	62.9	6.2	83.2%	48.0	39.3	11.4	165,397	330,051
DCA	3	44	110.6	39.8	80.6%	93.0	87.9	8.7	183,001	95,133
DEN	6	144	231.1	228.2	82.0%	315.5	179.0	21.1	384,384	183,132
DFW	7	174	247.9	84.8	80.5%	198.3	190.7	28.4	486,401	229,598
DTW	6	139	181.7	87.7	77.9%	104.1	113.5	17.8	315,031	206,868
EWR	3	91	369.6	58.2	69.5%	443.6	221.4	16.6	265,300	175,041
FLL	3	57	96.2	39.4	73.4%	44.5	99.7	11.0	180,546	150,147
HNL	4	47	88.0	32.1	90.8%	77.8	92.4	10.2	183,510	144,803
IAD	3	120	171.4	70.1	79.7%	142.5	153.8	13.2	154,286	396,566
IAH	5	82	172.6	59.9	83.2%	206.4	86.4	19.1	268,715	294,729
JFK	4	172	584.1	56.5	73.5%	570.9	245.0	20.3	296,228	66,452
LAS	4	95	134.3	68.7	77.2%	112.8	151.4	21.9	393,137	221,175
LAX	4	106	408.0	20.3	81.7%	211.2	271.5	30.3	454,920	195,526
LGA	2	74	216.1	18.2	72.3%	174.9	107.3	13.0	211,334	194,179
MCO	4	96	162.6	71.6	78.8%	90.5	175.5	16.8	255,548	104,049
MDW	5	29	88.2	48.0	79.8%	45.2	47.0	8.8	185,616	104,975
MEM	4	81	41.2	34.1	81.4%	76.4	31.2	5.7	223,707	170,907
MIA	4	107	343.6	113.5	75.1%	348.0	154.0	15.2	294,465	86,841
MSP	4	125	113.2	85.4	78.8%	104.6	123.9	18.2	338,496	193,451
ORD	6	171	352.0	199.9	74.3%	332.0	200.9	37.1	620,875	351,371
PDX	3	68	76.5	33.5	80.3%	94.7	65.7	6.9	135,552	123,777

Airport	Run-ways	Gates	Operating expense (US$ million)	Non-operating expense (US$ million)	On-time perfor-mance	Aero-nautical revenues (US$ million)	Non-aero-nautical revenues (US$ million)	Enplane-ments (million pax)	Air carrier operations	Other operations
PHL	4	63	147.7	60.4	71.8%	143.4	69.2	15.7	291,731	244,421
PHX	3	105	143.0	54.1	80.4%	90.7	128.4	21.2	409,711	140,468
PIT	4	50	72.9	34.9	80.8%	77.4	47.7	5.3	82,491	186,132
SAN	1	45	91.4	13.4	81.3%	50.2	57.9	8.8	151,925	73,498
SEA	2	96	155.0	120.9	76.0%	173.8	138.0	14.7	254,829	86,641
SFO	4	67	276.6	259.7	78.6%	296.6	180.9	16.4	241,492	112,104
SLC	4	83	65.0	1.3	84.3%	40.7	50.5	10.9	168,433	281,782
STL	6	87	79.5	37.5	81.7%	70.8	40.2	7.2	138,320	158,099
TPA	3	60	73.1	26.1	79.1%	43.0	101.9	9.5	158,712	110,435
2006										
ATL	5	184	73.4	37.5	71.8%	53.2	101.1	41.6	673,734	302,713
BOS	6	102	234.0	67.5	74.6%	213.9	186.5	13.6	212,509	203,060
BWI	4	78	143.3	24.1	79.3%	69.7	62.9	10.8	209,198	92,849
CLE	4	68	62.4	45.7	80.5%	63.0	42.8	5.5	73,733	175,733
CLT	3	85	59.3	44.0	76.8%	68.4	54.7	15.0	259,276	251,642
CVG	4	87	58.2	11.3	83.8%	41.6	39.6	8.1	93,312	252,446
DCA	3	44	109.0	38.5	79.2%	96.0	88.8	9.1	157,864	120,607
DEN	6	144	256.1	222.2	78.4%	302.5	198.4	23.1	428,653	180,861
DFW	7	174	334.9	234.3	77.3%	259.8	225.7	29.0	481,026	221,660
DTW	6	139	179.7	102.3	76.9%	92.0	125.9	17.7	289,637	192,103
EWR	3	91	372.6	55.3	67.2%	434.3	231.1	17.9	273,776	174,786
FLL	3	57	109.8	43.0	79.1%	54.8	105.3	10.6	179,848	117,227
HNL	4	47	103.6	29.7	91.1%	77.3	93.2	10.0	187,058	127,980
IAD	3	120	180.9	75.1	75.2%	137.5	148.9	11.2	163,072	256,944
IAH	5	82	160.0	77.8	78.7%	230.7	100.3	20.6	271,072	332,041
JFK	4	172	576.2	57.4	72.3%	553.8	244.4	21.3	319,380	77,354
LAS	4	95	149.4	83.5	75.9%	121.5	170.5	22.6	402,222	217,252
LAX	4	106	451.3	19.1	78.6%	259.8	280.9	30.2	463,341	193,345
LGA	2	74	216.6	18.4	69.5%	168.6	112.9	12.9	206,305	199,906
MCO	4	96	169.8	68.8	80.5%	92.2	188.7	17.0	265,134	90,878
MDW	5	29	97.4	52.8	75.1%	56.7	48.9	9.4	193,307	104,993
MEM	4	81	45.8	34.8	78.9%	80.9	29.1	5.6	212,571	175,322
MIA	4	107	348.1	113.3	77.3%	378.3	146.9	15.8	297,032	89,412
MSP	4	125	111.6	94.9	80.5%	87.4	136.7	17.4	297,424	178,209
ORD	6	171	354.6	201.0	68.5%	340.3	205.7	37.2	629,241	329,402
PDX	3	68	77.5	34.0	80.9%	91.6	71.0	7.1	139,419	119,968
PHL	4	63	160.2	51.7	71.3%	158.4	74.2	15.7	271,341	244,527
PHX	3	105	163.7	43.4	79.8%	90.8	155.0	21.6	411,928	122,123
PIT	4	50	73.5	39.7	77.9%	83.1	45.3	5.1	89,223	145,315
SAN	1	45	101.4	9.4	80.2%	53.0	66.5	8.9	156,335	69,449
SEA	2	96	160.8	130.5	76.2%	187.6	149.9	15.0	253,507	85,917
SFO	4	67	292.3	201.7	73.1%	259.0	192.5	16.6	248,297	110,905

Airport	Run-ways	Gates	Operating expense (US$ million)	Non-operating expense (US$ million)	On-time perfor-mance	Aero-nautical revenues (US$ million)	Non-aero-nautical revenues (US$ million)	Enplane-ments (million pax)	Air carrier operations	Other operations
SLC	4	83	68.7	2.2	84.3%	41.7	55.1	10.5	165,035	252,257
STL	6	87	82.1	48.9	79.3%	77.5	38.3	7.5	136,131	144,981
TPA	3	60	82.6	31.4	80.5%	47.1	109.1	9.4	161,690	95,247
2007										
ATL	5	184	148.3	140.8	73.2%	137.2	215.4	43.5	722,461	269,156
BOS	6	102	244.1	75.8	72.5%	220.4	199.7	13.8	205,620	197,164
BWI	4	78	157.4	24.1	77.1%	77.1	66.1	11.0	209,182	84,215
CLE	4	68	69.4	47.4	77.1%	61.4	44.4	5.6	72,118	173,052
CLT	3	85	62.6	45.9	71.6%	66.6	67.0	16.8	289,850	236,093
CVG	4	87	62.0	16.8	77.2%	44.3	40.7	7.8	90,151	238,110
DCA	3	44	126.4	45.3	74.4%	97.4	98.6	9.1	155,425	124,036
DEN	6	144	290.8	237.0	76.0%	316.3	213.8	24.5	451,192	168,736
DFW	7	174	342.2	233.1	71.4%	223.0	344.6	28.8	477,920	208,752
DTW	6	139	184.2	101.6	75.4%	95.4	133.7	17.7	271,034	196,200
EWR	3	91	398.2	56.6	63.6%	472.3	239.5	18.2	273,652	168,042
FLL	3	57	125.9	42.6	76.0%	57.8	113.5	11.1	193,712	114,230
HNL	4	47	116.2	26.9	91.4%	78.9	90.4	10.6	182,455	123,734
IAD	3	120	210.1	94.8	73.4%	150.9	163.6	12.0	185,807	232,250
IAH	5	82	168.0	91.4	79.2%	248.7	106.9	20.8	290,886	312,751
JFK	4	172	597.3	50.9	65.9%	612.6	260.2	23.6	356,364	100,471
LAS	4	95	202.7	80.2	76.5%	145.8	184.5	23.1	407,618	211,669
LAX	4	106	468.7	15.4	78.3%	329.1	232.0	30.9	467,193	213,761
LGA	2	74	227.9	16.4	65.0%	185.3	109.9	12.5	200,814	196,466
MCO	4	96	180.3	68.2	77.9%	97.8	201.2	17.8	291,400	76,460
MDW	5	29	111.3	57.4	75.8%	54.9	53.2	9.7	198,949	105,450
MEM	4	81	50.4	33.3	79.0%	78.5	33.0	5.7	212,347	164,181
MIA	4	107	357.4	123.4	72.1%	387.2	167.8	16.4	294,307	91,993
MSP	4	125	124.6	95.6	75.0%	95.9	142.3	17.1	286,310	167,256
ORD	6	171	398.1	206.5	66.2%	429.4	226.1	36.9	617,135	309,838
PDX	3	68	82.6	29.8	79.9%	89.2	77.4	7.4	148,756	115,258
PHL	4	63	179.2	55.1	68.1%	165.0	80.9	15.9	274,720	224,963
PHX	3	105	186.6	42.7	77.7%	94.0	179.4	21.7	408,641	121,108
PIT	4	50	79.3	31.0	74.0%	90.9	50.2	5.0	87,354	123,365
SAN	1	45	104.6	16.6	80.5%	56.7	68.7	9.4	161,896	72,474
SEA	2	96	171.6	124.8	74.0%	195.0	152.5	15.7	276,954	69,198
SFO	4	67	311.6	197.8	72.7%	292.4	211.6	17.6	262,135	117,365
SLC	4	83	70.5	2.6	82.0%	45.1	60.3	10.8	166,816	254,171
STL	6	87	94.6	124.5	76.4%	88.3	37.5	7.6	128,372	127,586
TPA	3	60	89.7	32.8	78.8%	47.8	118.3	9.5	169,973	88,342
2008										
ATL	5	184	175.7	112.3	76.2%	161.4	224.8	44.0	750,597	227,487
BOS	6	102	261.5	78.6	76.4%	229.3	220.5	12.8	193,229	182,161

Airport	Run-ways	Gates	Operating expense (US$ million)	Non-operating expense (US$ million)	On-time perfor-mance	Aero-nautical revenues (US$ million)	Non-aero-nautical revenues (US$ million)	Enplane-ments (million pax)	Air carrier operations	Other operations
BWI	4	78	154.4	23.1	80.2%	92.3	78.1	10.7	204,221	68,295
CLE	4	68	74.9	43.2	78.6%	62.9	48.5	5.4	67,842	168,127
CLT	3	85	70.2	54.1	79.1%	71.4	75.8	17.5	315,130	222,468
CVG	4	87	62.9	14.5	79.5%	46.2	38.5	6.7	78,604	207,464
DCA	3	44	128.2	105.2	80.4%	97.3	100.8	8.8	172,122	105,799
DEN	6	144	373.8	258.6	78.5%	319.6	222.1	24.8	460,311	165,533
DFW	7	174	371.3	251.2	75.0%	232.9	394.2	27.5	475,921	179,385
DTW	6	139	199.1	96.7	80.1%	121.0	134.6	17.1	241,757	222,027
EWR	3	91	415.7	100.5	65.7%	483.4	235.0	17.7	287,967	154,130
FLL	3	57	123.3	42.1	77.1%	63.0	124.2	11.1	194,695	100,970
HNL	4	47	137.4	37.8	88.1%	82.0	92.4	9.1	153,256	127,714
IAD	3	120	197.9	268.1	75.0%	185.5	167.4	11.6	186,097	205,529
IAH	5	82	181.0	85.4	78.7%	256.9	117.7	20.1	276,828	301,460
JFK	4	172	657.5	85.9	72.1%	673.1	277.5	23.8	356,397	90,571
LAS	4	95	239.4	166.7	77.9%	163.9	264.1	21.6	388,750	190,196
LAX	4	106	517.9	18.3	78.9%	432.3	280.3	29.5	453,232	169,274
LGA	2	74	237.4	25.5	68.9%	205.8	101.9	11.6	204,053	180,027
MCO	4	96	198.1	64.2	79.0%	134.7	208.2	17.5	293,229	50,171
MDW	5	29	109.7	62.0	78.7%	70.0	55.0	8.7	188,477	79,043
MEM	4	81	50.6	32.1	80.6%	78.7	31.5	5.5	208,188	154,790
MIA	4	107	382.9	154.6	71.3%	387.9	174.0	16.5	297,779	73,740
MSP	4	125	126.7	86.7	80.5%	98.5	143.0	16.5	286,192	163,780
ORD	6	171	428.5	224.2	68.1%	461.8	222.5	34.0	580,967	300,599
PDX	3	68	87.4	31.2	80.8%	96.8	81.1	7.2	155,088	96,965
PHL	4	63	189.0	52.1	75.0%	160.1	90.5	15.8	278,807	213,231
PHX	3	105	196.6	41.1	81.1%	97.4	184.5	20.4	391,518	109,007
PIT	4	50	82.1	28.6	78.2%	81.1	48.6	4.4	88,201	79,165
SAN	1	45	114.0	17.8	79.8%	59.5	76.2	9.2	164,382	61,775
SEA	2	96	195.2	137.6	78.2%	204.4	154.0	16.0	306,425	38,632
SFO	4	67	326.1	208.6	72.1%	300.7	235.0	18.3	284,350	103,620
SLC	4	83	77.1	7.9	85.0%	45.4	65.3	10.3	169,154	220,303
STL	6	87	99.4	75.5	78.2%	91.5	43.1	7.2	124,898	122,741
TPA	3	60	91.8	31.2	79.7%	48.3	119.1	9.1	165,448	72,437

E.3.1.2 Operational Efficiency Benchmark

Table E.6 - Case study 3 benchmark data for operational efficiency study

Airport	Total runway capacity (number of flights per 15 minutes)	Total passenger delay (millions of hours)	Enplanements (million pax)	Total operations (thousands of movements)
2005				
ATL	52	19.14	42.62	980.4
BOS	25	4.74	13.23	421.4
BWI	19	2.31	10.24	303.9
CLE	23	1.29	5.59	259.7
CLT	30	2.38	14.34	523.2
CVG	47	3.34	11.37	495.4
DCA	17	2.38	8.74	278.1
DEN	58	4.09	21.11	567.5
DFW	50	6.69	28.38	716.0
DTW	37	3.96	17.78	521.9
EWR	21	6.38	16.55	440.3
FLL	20	3.80	10.96	330.7
HNL	28	0.97	10.21	328.3
IAD	32	2.53	13.15	550.9
IAH	41	4.40	19.15	563.4
JFK	20	4.68	20.34	362.7
LAS	27	4.43	21.86	614.3
LAX	37	4.73	30.26	650.4
LGA	18	5.90	13.03	405.5
MCO	40	3.67	16.77	359.6
MDW	17	1.99	8.83	290.6
MEM	35	1.08	5.73	394.6
MIA	32	3.20	15.24	381.3
MSP	35	4.00	18.16	531.9
ORD	46	10.05	37.10	972.2
PDX	27	1.38	6.91	259.3
PHL	24	5.06	15.72	536.2
PHX	34	3.89	21.25	550.2
PIT	38	1.18	5.28	268.6
SAN	12	1.79	8.80	225.4
SEA	22	3.62	14.67	341.5
SFO	22	3.54	16.41	353.6
SLC	34	2.36	10.85	450.2
STL	30	1.17	7.23	296.4
TPA	27	2.29	9.46	269.1
2006				
ATL	52	12.24	41.59	976.4
BOS	25	4.21	13.58	415.6

Airport	Total runway capacity (number of flights per 15 minutes)	Total passenger delay (millions of hours)	Enplanements (million pax)	Total operations (thousands of movements)
BWI	19	2.30	10.84	302.0
CLE	23	1.30	5.50	249.5
CLT	30	2.95	15.02	510.9
CVG	47	1.90	8.09	345.8
DCA	17	2.34	9.08	278.5
DEN	58	6.83	23.12	609.5
DFW	50	7.97	29.01	702.7
DTW	37	3.78	17.70	481.7
EWR	21	7.07	17.90	448.6
FLL	20	2.17	10.56	297.1
HNL	28	1.19	9.99	315.0
IAD	32	2.68	11.21	420.0
IAH	41	4.41	20.59	603.1
JFK	20	4.72	21.33	396.7
LAS	27	4.02	22.61	619.5
LAX	37	5.66	30.19	656.7
LGA	18	5.74	12.95	406.2
MCO	40	3.31	17.01	356.0
MDW	17	2.36	9.42	298.3
MEM	35	1.06	5.61	387.9
MIA	32	2.52	15.84	386.4
MSP	35	3.43	17.38	475.6
ORD	46	16.20	37.25	958.6
PDX	27	1.36	7.07	259.4
PHL	24	4.25	15.71	515.9
PHX	34	3.37	21.58	534.1
PIT	38	1.01	5.05	234.5
SAN	12	2.08	8.86	225.8
SEA	22	3.56	15.02	339.4
SFO	22	4.04	16.56	359.2
SLC	34	2.33	10.54	417.3
STL	30	1.58	7.48	281.1
TPA	27	1.97	9.39	256.9
2007				
ATL	52	13.99	43.48	991.6
BOS	25	6.69	13.81	402.8
BWI	19	3.46	11.02	293.4
CLE	23	1.65	5.63	245.2
CLT	30	5.06	16.79	525.9
CVG	47	2.52	7.81	328.3
DCA	17	4.07	9.11	279.5
DEN	58	7.85	24.50	619.9
DFW	50	17.08	28.84	686.7
DTW	37	7.08	17.68	467.2

Airport	Total runway capacity (number of flights per 15 minutes)	Total passenger delay (millions of hours)	Enplanements (million pax)	Total operations (thousands of movements)
EWR	21	9.56	18.24	441.7
FLL	20	2.86	11.13	307.9
HNL	28	1.50	10.59	306.2
IAD	32	4.11	11.97	418.1
IAH	41	4.61	20.82	603.6
JFK	20	9.08	23.62	456.8
LAS	27	4.92	23.06	619.3
LAX	37	7.56	30.85	681.0
LGA	18	9.20	12.54	397.3
MCO	40	4.81	17.80	367.9
MDW	17	2.71	9.66	304.4
MEM	35	2.24	5.66	376.5
MIA	32	3.55	16.40	386.3
MSP	35	7.00	17.11	453.6
ORD	46	22.23	36.87	927.0
PDX	27	1.60	7.38	264.0
PHL	24	6.21	15.88	499.7
PHX	34	5.17	21.74	529.7
PIT	38	1.62	4.98	210.7
SAN	12	2.50	9.35	234.4
SEA	22	4.70	15.70	346.2
SFO	22	5.72	17.59	379.5
SLC	34	2.76	10.83	421.0
STL	30	2.45	7.61	256.0
TPA	27	2.67	9.53	258.3
2008				
ATL	52	10.74	43.95	978.1
BOS	25	4.13	12.85	375.4
BWI	19	2.25	10.66	272.5
CLE	23	1.35	5.42	236.0
CLT	30	3.57	17.49	537.6
CVG	47	1.63	6.74	286.1
DCA	17	2.14	8.76	277.9
DEN	58	5.57	24.84	625.8
DFW	50	10.46	27.47	655.3
DTW	37	3.91	17.13	463.8
EWR	21	6.89	17.69	442.1
FLL	20	2.26	11.09	295.7
HNL	28	1.27	9.11	281.0
IAD	32	2.46	11.56	391.6
IAH	41	4.89	20.10	578.3
JFK	20	6.41	23.79	447.0
LAS	27	4.76	21.60	578.9
LAX	37	5.76	29.52	622.5

Airport	Total runway capacity (number of flights per 15 minutes)	Total passenger delay (millions of hours)	Enplanements (million pax)	Total operations (thousands of movements)
LGA	18	5.55	11.58	384.1
MCO	40	4.16	17.49	343.4
MDW	17	1.98	8.67	267.5
MEM	35	1.56	5.46	363.0
MIA	32	3.26	16.48	371.5
MSP	35	3.65	16.50	450.0
ORD	46	15.53	33.96	881.6
PDX	27	1.60	7.17	252.1
PHL	24	3.79	15.79	492.0
PHX	34	4.57	20.41	500.5
PIT	38	0.93	4.36	167.4
SAN	12	2.19	9.22	226.2
SEA	22	3.64	16.02	345.1
SFO	22	5.19	18.35	388.0
SLC	34	2.13	10.25	389.5
STL	30	1.91	7.20	247.6
TPA	27	2.20	9.09	237.9

E.3.1.3 Investment Quality Benchmark

Table E.7 - Case study 3 benchmark data for investment quality study

Airport	Regional growth		Air service				Financial factors	
	Regional population growth (2007-2008)	Regional GDP growth (2007-2008)	Growth in enplaned pax (2007-2008)	O&D pax as portion of all pax (2008)	Portion of the main carrier's pax enplaned at this airport (2008)	Portion of total OEP-35 pax enplaned at this airport (2008)	Debt service coverage ratio (2008)	Non-aero-nautical revenue as portion of total revenue (2008)
ATL	2.19%	0.94%	1.07%	36.19%	43.89%	7.43%	2.81	58.20%
BOS	0.46%	3.36%	-6.96%	96.43%	12.68%	2.17%	1.26	49.02%
BWI	0.74%	3.99%	-3.30%	80.26%	9.90%	1.80%	1.62	45.85%
CLE	-0.31%	1.43%	-3.71%	72.00%	4.68%	0.92%	1.20	43.52%
CLT	3.36%	1.59%	4.18%	26.32%	22.43%	2.96%	1.63	51.49%
CVG	0.30%	2.04%	-13.76%	28.95%	4.18%	1.14%	0.71	45.44%
DCA	0.74%	3.99%	-3.83%	77.46%	4.77%	1.48%	1.11	50.87%
DEN	2.17%	4.79%	1.38%	53.13%	18.14%	4.20%	0.62	41.00%
DFW	2.38%	4.91%	-4.76%	45.09%	29.01%	4.64%	1.52	62.86%
DTW	-0.73%	0.06%	-3.12%	55.71%	27.47%	2.89%	0.86	52.66%
EWR	0.45%	4.54%	-3.05%	86.17%	27.58%	2.99%	3.70	32.71%
FLL	0.42%	0.47%	-0.33%	95.63%	44.42%	1.87%	1.09	66.34%
HNL	0.50%	3.75%	-14.03%	79.65%	77.60%	1.54%	1.14	53.00%
IAD	0.74%	3.99%	-3.46%	64.46%	8.99%	1.95%	1.05	47.43%
IAH	2.33%	7.39%	-3.49%	44.13%	34.95%	3.40%	1.10	31.43%
JFK	0.45%	4.54%	0.74%	82.85%	40.94%	4.02%	4.52	29.19%
LAS	2.08%	1.37%	-6.34%	81.46%	14.32%	3.65%	0.25	61.70%
LAX	0.82%	2.37%	-4.31%	82.25%	6.35%	4.99%	4.31	39.33%
LGA	0.45%	4.54%	-7.63%	92.14%	3.25%	1.96%	4.58	33.11%
MCO	1.28%	1.83%	-1.73%	93.72%	6.70%	2.96%	0.53	60.71%
MDW	0.77%	1.96%	-10.24%	70.04%	12.78%	1.46%	0.56	44.02%
MEM	0.52%	1.39%	-3.54%	37.76%	6.48%	0.92%	1.15	28.58%
MIA	0.42%	0.47%	0.47%	80.77%	15.50%	2.79%	1.21	30.97%
MSP	1.01%	3.86%	-3.54%	52.11%	28.38%	2.79%	0.69	59.20%
ORD	0.77%	1.96%	-7.88%	53.91%	22.73%	5.74%	1.13	32.51%
PDX	1.89%	2.54%	-2.73%	85.81%	2.25%	1.21%	1.79	45.61%
PHL	0.26%	2.97%	-0.56%	64.82%	13.67%	2.67%	1.83	36.12%
PHX	2.78%	0.46%	-6.12%	60.80%	17.32%	3.45%	4.28	65.44%
PIT	-0.13%	3.82%	-12.38%	92.71%	1.86%	0.74%	0.62	37.47%
SAN	1.40%	4.45%	-1.36%	94.39%	5.93%	1.56%	10.74	56.18%
SEA	1.41%	4.00%	2.05%	77.93%	48.87%	2.71%	1.69	42.97%
SFO	1.52%	2.09%	4.30%	82.20%	13.30%	3.10%	0.21	43.87%
SLC	1.86%	3.19%	-5.34%	56.07%	7.20%	1.73%	1.35	58.99%
STL	0.40%	5.22%	-5.31%	78.55%	4.25%	1.22%	1.08	32.04%
TPA	0.68%	-0.21%	-4.62%	92.60%	4.86%	1.54%	1.04	71.14%

E.3.2 Results

E.3.2.1 Original Study

Table E.8 - Case study 3 benchmark results for original study

Airport	Results			
	2005	2006	2007	2008
ATL	0.4380	0.6240	0.3397	0.2879
BOS	0.2366	0.2328	0.2255	0.2466
BWI	0.3264	0.3237	0.3138	0.3384
CLE	0.5526	0.6485	0.6129	0.6023
CLT	0.7465	0.6512	0.6306	0.6463
CVG	0.6007	0.7238	0.6866	0.7256
DCA	0.5856	0.5730	0.5366	0.6012
DEN	0.1819	0.1733	0.1676	0.1794
DFW	0.1478	0.1414	0.1667	0.1667
DTW	0.1948	0.2152	0.2257	0.2310
EWR	0.2719	0.2743	0.2570	0.2484
FLL	0.4116	0.4417	0.4229	0.4454
HNL	0.6170	0.6170	0.6170	0.6170
IAD	0.3333	0.2750	0.2677	0.2837
IAH	0.3244	0.3537	0.3537	0.3537
JFK	0.2500	0.2500	0.2500	0.2500
LAS	0.2611	0.2552	0.2557	0.2699
LAX	0.2736	0.2736	0.2344	0.3846
LGA	0.3982	0.3812	0.3558	0.3909
MCO	0.2624	0.2668	0.2575	0.2709
MDW	0.8793	0.8241	0.8293	0.8935
MEM	0.8965	0.8662	0.8646	0.9145
MIA	0.2242	0.2299	0.2139	0.2193
MSP	0.3163	0.3623	0.3318	0.3645
ORD	0.1521	0.1682	0.1680	0.1691
PDX	0.4766	0.5247	0.5336	0.5309
PHL	0.3643	0.3603	0.3429	0.3916
PHX	0.2952	0.2919	0.2836	0.3066
PIT	0.5162	0.5331	0.5143	0.5470
SAN	0.8957	0.8802	0.8805	0.9056
SEA	0.4186	0.4182	0.4048	0.4437
SFO	0.3750	0.3473	0.3444	0.3543
SLC	0.9286	0.9249	0.8970	0.9645
STL	0.4663	0.4854	0.4454	0.4516
TPA	0.4915	0.4899	0.4846	0.4986

E.3.2.2 Operational Efficiency Benchmark

Table E.9 - Case study 3 benchmark results for operational efficiency study

Airport	Results			
	2005	2006	2007	2008
ATL	1.0000	1.0000	1.0000	1.0000
BOS	1.3474	1.3729	1.4223	1.3891
BWI	1.1828	1.1039	1.3561	1.0330
CLE	1.1697	1.0000	1.0000	1.0000
CLT	1.0000	1.0485	1.1882	1.0000
CVG	1.2033	1.3081	1.2315	1.2064
DCA	1.2083	1.1313	1.3066	1.0076
DEN	1.1026	1.2164	1.1928	1.0378
DFW	1.0805	1.1174	1.4127	1.4438
DTW	1.1932	1.2317	1.5150	1.1535
EWR	1.0350	1.0211	1.0868	1.0537
FLL	1.3237	1.1112	1.1421	1.0075
HNL	1.0000	1.0000	1.0000	1.0000
IAD	1.0000	1.2261	1.3136	1.0833
IAH	1.1512	1.0549	1.0000	1.0702
JFK	1.0000	1.0000	1.0000	1.0000
LAS	1.0000	1.0000	1.0000	1.0000
LAX	1.0000	1.0000	1.0000	1.0000
LGA	1.0000	1.0000	1.0099	1.0200
MCO	1.4668	1.2504	1.2733	1.2714
MDW	1.0686	1.0635	1.0052	1.0000
MEM	1.0000	1.0000	1.0022	1.0000
MIA	1.3779	1.0773	1.1008	1.1004
MSP	1.1738	1.1746	1.5541	1.1345
ORD	1.0000	1.0000	1.0000	1.0008
PDX	1.4080	1.2494	1.1566	1.2183
PHL	1.0160	1.0578	1.0969	1.0000
PHX	1.1041	1.0000	1.0954	1.1354
PIT	1.4561	1.0000	1.5099	1.0000
SAN	1.0000	1.0000	1.0000	1.0000
SEA	1.2594	1.2410	1.2266	1.0860
SFO	1.1371	1.2172	1.1716	1.1613
SLC	1.1745	1.1682	1.0146	1.0181
STL	1.2433	1.3659	1.5464	1.4612
TPA	1.5890	1.4098	1.5389	1.3942

E.3.2.3 Investment Quality Benchmark

Table E.10 - Case study 3 benchmark results for investment quality study

Airport	Results		
	Regional growth	Air service	Financial factors
ATL	0.0000	0.2784	0.3730
BOS	0.0000	0.4413	0.5476
BWI	0.3728	0.4695	0.5347
CLE	0.4162	0.4861	0.5726
CLT	0.0000	0.4128	0.0000
CVG	0.5122	0.4726	0.5537
DCA	0.3998	0.4207	0.5347
DEN	0.2341	0.5012	0.3769
DFW	0.3396	0.3034	0.2914
DTW	0.3410	0.3968	0.5821
EWR	0.2043	0.5369	0.5482
FLL	0.0000	0.3274	0.5492
HNL	0.0000	0.3918	0.5459
IAD	0.3921	0.4563	0.5347
IAH	0.3265	0.5411	0.0000
JFK	0.0000	0.5478	0.5482
LAS	0.2638	0.3582	0.3963
LAX	0.0000	0.5099	0.5309
LGA	0.3796	0.5355	0.5482
MCO	0.0000	0.3499	0.5009
MDW	0.4416	0.4828	0.5335
MEM	0.4552	0.5495	0.5452
MIA	0.2231	0.5425	0.5492
MSP	0.3522	0.3450	0.5198
ORD	0.2648	0.5376	0.5335
PDX	0.3864	0.4714	0.4317
PHL	0.3267	0.5243	0.5552
PHX	0.3627	0.0414	0.3876
PIT	0.4481	0.5186	0.5676
SAN	0.2765	0.0000	0.4908
SEA	0.0000	0.4897	0.4894
SFO	0.0000	0.4838	0.4790
SLC	0.4352	0.3200	0.4372
STL	0.4254	0.5391	0.5499
TPA	0.3748	0.0000	0.5377

References

References

Abbott, M. & Wu, S., 2002. Total Factor Productivity and Efficiency of Australian Airports. *Australian Economic Review*, 35, 244-260.

Adler, N. & Berechman, J., 2001. Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis. *Transport Policy*, 8(3), 171-181.

Ahmed, P.K. & Rafiq, M., 1998. Integrated benchmarking: a holistic examination of select techniques for benchmarking analysis. *Benchmarking: An International Journal*, 5, 225-242.

Air Transport Research Society, 2009. 2009 Airport Benchmarking Report Information Package. Available at: http://www.atrsworld.org/docs/Information%20Package%20_2009%20-%20Report_.pdf [Accessed December 4, 2009].

Airbus S.A.S., 2010. Summary historical orders and deliveries 1974-2009. Available at: http://www.airbus.com/fileadmin/media_gallery/files/reports_results_reviews/Summary_Historial_Orders_Deliveries_1974-2009.xls [Accessed March 22, 2010].

Airports Council International, 2009. Airport Service Benchmarking. Available at: http://www.airportservicequality.aero/index.html [Accessed December 4, 2009].

Airports Council International - North America, 2009a. *Airport Capital Development Costs 2009-2013*, Airports Council International - North America. Available at: http://www.aci-na.org/static/entransit/CapitalNeedsSurveyReport2009.pdf [Accessed December 10, 2009].

Airports Council International - North America, 2009b. Airports Council International - North America - Benchmarking. Available at: http://aci-na.org/index/toolbox_benchmarking_main.

Ali, A.I. & Seiford, L.M., 1993. Computational accuracy and infinitesimals in data envelopment analysis. *INFOR*, 31(4), 290.

Allen, R. et al., 1997. Weights restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research*, 73(0), 13-34.

Alodhaibi, K., Brodsky, A. & Mihaila, G.A., 2010. COD: Iterative Utility Elicitation for Diversified Composite Recommendations. In *2010 43rd Hawaii International Conference on System Sciences*. 2010 43rd Hawaii International Conference on System Sciences (HICSS). Honolulu, Hawaii, USA, pp. 1-10.

American Airlines, 2010. Basic Measurements in the Airline Business. *Basic Measurements in the Airline Business*. Available at: http://www.aa.com/i18n/amrcorp/corporateInformation/facts/measurements.jsp [Accessed April 15, 2010].

American Association of Airport Executives, 2006. AAAE Rates and Charges Survey. Available at: http://www.aaae.org/federal_affairs/regulatory_affairs/rates_and_charges_survey/ [Accessed December 4, 2009].

American University School of Communication, 2010. *Flying Cheap*, Available at: http://investigativereportingworkshop.org/flying-cheap/airport/MHT/ [Accessed March 22, 2010].

Andersen, P. & Petersen, N.C., 1993. A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, 39(10), 1261-1264.

Association of Metropolitan Planning Organizations, 2010. What is AMPO. Available at: [Accessed February 5, 2010].

A-Z World Airports, 2010. A-Z World Airports. *A-Z World Airports*. Available at: http://www.azworldairports.com/cfm/homepage.cfm [Accessed May 28, 2010].

Banker, R.D., Charnes, A. & Cooper, W.W., 1984. Some Models for Estimating

Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078-1092.

Banker, R.D., 1996. Hypothesis Tests Using Data Envelopment Analysis. *Journal of Productivity Analysis*, 7(2/3), 139-159.

Banxia, 2010. Banxia Frontier Analyst. Available at: http://www.banxia.com/frontier/ [Accessed June 1, 2010].

Barros, C.P., 2008. Airports in Argentina: Technical efficiency in the context of an economic crisis. *Journal of Air Transport Management*, 14(6), 315-319.

Barros, C.P. & Dieke, P., 2007. Performance evaluation of Italian airports: A data envelopment analysis. *Journal of Air Transport Management*, 13(4), 184-191.

Bazargan, M. & Vasigh, B., 2003. Size versus efficiency: a case study of US commercial airports. *Journal of Air Transport Management*, 9(4), 187-193.

Belobaba, P., Odoni, A.R. & Barnhart, C., 2009. *The Global Airline Industry*, American Institute of Aeronautics & Astronautics.

Beretta, S., Dossi, A. & Grove, H., 1998. Methodological strategies for benchmarking accounting processes. *Benchmarking: An International Journal*, 5, 165-183.

Bombardier, 2010. Bombardier CRJ Series. *Bombardier CRJ Series*. Available at: http://www.crj.bombardier.com/CRJ/en/home.jsp [Accessed March 22, 2010].

Bureau of Economic Analysis, U.S. Department of Commerce, 2010. BEA: Gross Domestic Product by Metropolitan Area. Available at: http://www.bea.gov/regional/gdpmetro/ [Accessed March 17, 2010].

Bureau of Transportation Statistics, 2010a. Airline On-Time Tables. *Airline On-Time Tables*. Available at: http://www.bts.gov/programs/airline_information/airline_ontime_tables/ [Accessed March 26, 2010].

Bureau of Transportation Statistics, 2010b. BTS Form 41, Schedules T100 and T100(f) Air Carrier Data. Available at:

http://www.bts.gov/programs/statistical_policy_and_research/source_and_accuracy_compendium/form41_schedule.html [Accessed March 17, 2010].

Bureau of Transportation Statistics, 2010c. RITA | BTS | Transtats. *RITA | BTS | Transtats*. Available at: http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=125&DB_URL= [Accessed March 17, 2010].

Button, K.J. & Stough, R., 2000. *Air transport networks*, Edward Elgar Publishing.

Camp, R.C., 1993. A bible for benchmarking, by Xerox. *Financial Executive*, 9(4), 23-28.

Camp, R.C., 1989. Benchmarking: the search for industry best practices that lead to superior performance. *Quality Progress*, (1-4).

Carney, M. & Mew, K., 2003. Airport governance reform: a strategic management perspective. *Journal of Air Transport Management*, 9(4), 221-232.

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

Charnes, A., Cooper, W.W. & Rhodes, E., 1979. Short Communication: Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 3(4), 339.

Charnes, A., 1994. *Data envelopment analysis*, Springer.

Chellie Cameron, MWAA, 2009. Telephone call betwee Chellie Cameron, MWAA, and David Schaar, George Mason University.

Chen, Y. & Iqbal Ali, A., 2004. DEA Malmquist productivity measure: New insights with an application to computer industry. *European Journal of Operational Research*, 159(1), 239-249.

City of Cleveland, Department of Port Control, 2007. Strategic Plan 2007-2010. Available at: http://www.clevelandairport.com/Portals/Documents/2007-2010%20Strategic%20Plan.pdf [Accessed January 27, 2010].

Coelli, T., 2010. DEAP. *DEAP*. Available at: http://www.uq.edu.au/economics/cepa/deap.htm [Accessed June 1, 2010].

Cooper, R., 1990. Airports and economic development: An overview. *Transportation Research Record*, (1274), 125-133.

Cooper, W.W., Seiford, L.M. & Tone, K., 2006. *Introduction to data envelopment analysis and its uses*, Birkhäuser.

Cooper, W.W., Seiford, L.M. & Tone, K., 2006. *Data envelopment analysis*, Springer.

Customs and Border Protection, 2009. This is CBP. Available at: http://www.cbp.gov/xp/cgov/about/mission/cbp_is.xml [Accessed December 10, 2009].

Dallas Dawson, Tampa International Airport, 2009. Telephone call between Dallas Dawson, Tampa International Airport, and David Schaar, George Mason University.

Dattakumar, R. & Jagadeesh, R., 2003. A review of literature on benchmarking. *Benchmarking: An International Journal*, 10, 176-209.

DEA Frontier, 2010. DEA Frontier Software. Available at: http://www.deafrontier.net/index.html [Accessed June 1, 2010].

Denver International Airport, 2009. *Together We Soar: A Strategic Plan for Denver International Airport*,

DFW Airport, 2008. Building a Future Together: DFW International Airport Strategic Plan 2008. Available at: http://dfwairport.com/about/pdf/publications/14816_DFWAIR_STRATEGIC_PLAN_012508_resize.pdf [Accessed January 27, 2010].

DFW Airport, 2009. DFW Airport Administration. Available at: http://www.dfwairport.com/about/admin/index.php [Accessed December 9, 2009].

Dillingham, G.L., 2000. *Aviation and the Environment: FAA's Role in Major Airport Noise Programs*, DIANE Publishing.

Dillingham, G.L., 1996. *GAO Report on the Airport Improvement Program*, General Accounting Office; Resources, Community, and Economic Development Division. Available at: http://archive.gao.gov/papr2pdf/157264.pdf [Accessed December 10, 2009].

Doganis, R., 2002. *Flying off course*, Routledge.

Doganis, R., 1992. *The airport business*, Routledge.

Donohue, G.L., Shaver, R.D. & Edwards, E., 2008. *Terminal Chaos*, American Institute of Aeronautics & Astronautics.

Doyle, J. & Green, R., 1994. Efficiency and Cross-Efficiency in DEA: Derivations, Meanings and Uses. *The Journal of the Operational Research Society*, 45(5), 567-578.

Dyson, R.G. et al., 2001. Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2), 245-259.

Embraer, 2010. Centro Historico Embraer. *Centro Historico Embraer*. Available at: http://www.centrohistoricoembraer.com.br/en/historia-aeronaves/embraer170.htm [Accessed March 22, 2010].

Environmental Protection Agency, 2000. *Preliminary Data Summary: Airport Deicing Operations.*,

FAA, 2008. *National Plan of Integrated Airport Systems 2009-2013*, Available at: http://www.faa.gov/airports/planning_capacity/npias/reports/media/2009/npias_2009_narrative.pdf.

FAA, 2009. OEP Frequently Asked Questions - OEP 35 Airports. Available at: http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/oep/faq/Airports/index.cfm [Accessed December 3, 2009].

FAA Office of Aviation Policy and Plans, 2005. Weather Classification. Available at: http://aspm.faa.gov/aspm/Customer%20Satisfaction/Weather.pps.

FAA Office of Environment and Energy, 2005. *Aviation & Emissions: A Primer*, Available at: http://www.faa.gov/regulations_policies/policy_guidance/envir_policy/media/aeprimer.pdf [Accessed June 21, 2010].

Färe, R. & Knox Lovell, C.A., 1978. Measuring the technical efficiency of production. *Journal of Economic Theory*, 19(1), 150-162.

Federal Aviation Administration, 2010. Air Traffic Activity Systems (ATADS). *Air Traffic Activity Systems (ATADS)*. Available at: http://aspm.faa.gov/opsnet/sys/Airport.asp [Accessed May 28, 2010].

Federal Aviation Administration, 2004. *Airport Capacity Benchmark Report 2004*, Available at: http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/bench/2004download.htm.

Federal Aviation Administration, 2000. *Airports Capital Improvement Plan*, Available at: http://www.faa.gov/airports/resources/publications/orders/media/AIP_5100_39A.pdf [Accessed June 27, 2010].

Federal Aviation Administration, 2010a. APM Analysis: Definitions of Variables. *APM Analysis: Definitions of Variables*. Available at: http://aspmhelp.faa.gov/index.php/APM:_Definitions_of_Variables [Accessed June 4, 2010].

Federal Aviation Administration, 2010b. APM Taxi Times: Definitions of Variables. *APM Taxi Times: Definitions of Variables*. Available at: http://aspmhelp.faa.gov/index.php/APM_Taxi_Times:_Definitions_of_Variables [Accessed April 15, 2010].

Federal Aviation Administration, 2010c. Aviation Performance Metrics (APM). *Aviation Performance Metrics (APM)*. Available at: http://aspm.faa.gov/apm/sys/ [Accessed April 15, 2010].

Federal Aviation Administration, 2010d. Aviation System Performance Metrics. *Aviation System Performance Metrics*. Available at: http://aspm.faa.gov/ [Accessed April 8, 2010].

Federal Aviation Administration, 2007. *Capacity Needs in the National Airspace System*,

Federal Aviation Administration, 2010a. Compliance Activity Tracking System. Available at: http://cats.airports.faa.gov/ [Accessed February 4, 2010].

Federal Aviation Administration, 2010b. FAA Airport Improvement Program and American Recovery and REinvestment Act Projects - Fiscal Years 2005-2009. Available at: http://www.faa.gov/airports/aip/grant_histories/airport_projects/ [Accessed June 23, 2010].

Federal Aviation Administration, 2001. FAA Form 5100-127. Available at: http://forms.faa.gov/forms/faa5100-127.pdf [Accessed February 4, 2010].

Federal Aviation Administration, 2009. *The Economic Impact of Civil Aviation on the U.S. Economy*, Available at: http://www.faa.gov/air_traffic/publications/media/FAA_Economic_Impact_Rpt_2009.pdf [Accessed June 4, 2010].

Fitch Ratings, 2010. Fitch Ratings Definitions. Available at: http://www.fitchratings.com/creditdesk/public/ratings_defintions/index.cfm?rd_file=intro#cr_rtngs [Accessed June 16, 2010].

Forsgren, K., 2007. *Public Finance Criteria: Airport Revenue Bonds*, Standard & Poor's.

Frank Berardino, GRA Inc, 2009. Telephone call between Frank Berardino, GRA Inc, and David Schaar, George Mason University.

G. Hamzaee, R. & Vasigh, B., 2000. A simple revenue-cost perspective on US airport operations. *Journal of Air Transport Management*, 6(1), 61-64.

Gattoufi, S., Oral, M. & Reisman, A., 2004. A taxonomy for data envelopment analysis. *Socio-Economic Planning Sciences*, 38(2-3), 141-158.

Gillen, D. & Lall, A., 1997. Developing measures of airport productivity and performance: an application of data envelopment analysis. *Transportation Research Part E: Logistics and Transportation Review*, 33(4), 261-273.

Giorgetti, N., 2010. Matlab MEX Interfaces. *Matlab MEX Interfaces*. Available at: http://www.dii.unisi.it/~hybrid/tools/mex/downloads.html#cplexmex [Accessed June 5, 2010].

Graham, A., 2003. *Managing airports*, Butterworth-Heinemann.

Grossman, D., 2007. Dismantling Pittsburgh: Death of an airline hub. *USA Today*. Available at: http://www.usatoday.com/travel/columnist/grossman/2007-10-15-dismantling-pittsburgh-hub_N.htm [Accessed March 17, 2010].

Hansen, M. et al., 2009. Total Delay Impact Study.

Hazel, B., 2010. Personal conversation between David Schaar and Bob Hazel.

Hillsborough County Aviation Authority, 2006. Strategic Business Plan. Available at: http://www.tampaairport.com/airport_business/business_plan/2006_strategic_business_plan.pdf [Accessed January 27, 2010].

Hooper, P.G. & Hensher, D.A., 1997. Measuring total factor productivity of airports--an index number approach. *Transportation Research Part E: Logistics and Transportation Review*, 33(4), 249-259.

Hunt, E.E., 1944. *The power industry and the public interest*, The Twentieth Century Fund.

IBM, 2010. IBM ILOG CPLEX Optimizer. *IBM ILOG CPLEX Optimizer*. Available at: http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/ [Accessed June 1, 2010].

Intergovernmental Panel on Climate Change, 2000. *Aviation and the Global Atmosphere*, The Hague.

J.D. Power, 2008. 2008 North America Airport Satisfaction. Available at: http://www.jdpower.com/travel/articles/2008-North-America-Airport-Satisfaction [Accessed December 4, 2009].

J.D. Power, 2009. JDPower.com About Us. Available at: http://www.jdpower.com/about [Accessed December 4, 2009].

Jacobs Consultancy, 2009a. Jacobs Consultancy. Available at: http://www.jacobsconsultancy.co.uk/about_jacobs.shtml [Accessed December 4, 2009].

Jacobs Consultancy, 2009b. Publications. Available at: http://www.jacobsconsultancy.co.uk/publications.shtml#Aviation [Accessed December 4, 2009].

Jeff Gilley, National Business Aviation Association, 2009. Telephone call between Jeff Gilley, National Business Aviation Association, and David Schaar, George Mason University.

Jiang, J. et al., 2010. DEA efficiency measurement based on strongly efficient and strongly inefficient frontier.

Jim Wilding, formerly with MWAA, 2009. Telephone call between Jim Wilding, formerly with MWAA, and David Schaar, George Mason University.

Kindleberger, C.P., 1997. *Economic laws and economic history*, Cambridge University Press.

Kirk, R.S., 2003. *CRS Issue Brief for Congress: Airport Improvement Program*, Congressional Research Service.

Kleine, A., 2004. A general model framework for DEA. *Omega*, 32(1), 17-23.

Kleine, A., 2010. DEA Solver Online. *DEA Solver Online*. Available at: http://www.dea.uni-hohenheim.de/ [Accessed June 1, 2010].

Kruskal, W.H. & Wallis, W.A., 1952. Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583-621.

Kumar, V. & Sherry, L., 2009. Airport Throughput Capacity Limits for Demand Management Planning. In *2009 ICNS Conference*. 2009 ICNS Conference. Available at: http://catsr.ite.gmu.edu/pubs/ICNS_Kumar_Sherry.pdf [Accessed April 8, 2010].

Kurt Krummenacker, Moody's, 2009. Telephone call between Kurt Krummenacker, Moody's, and David Schaar, George Mason University.

Lee, J., 2008. Gigantic Airbus A380 Arrives, at a Diminished J.F.K. *The New York Times*. Available at: Gigantic Airbus A380 Arrives, at a Diminished J.F.K. [Accessed June 5, 2010].

Lehman, S., Wydick, E. & McDermott, M., 2009. U.S. Airports: Seeking Stability in an Uncertain Airline Market.

Lincoln Laboratory, Massachusetts Institute of Technology, 2010. Integrated Terminal Weather System. Available at: http://www.ll.mit.edu/mission/aviation/faawxsystems/itws.html [Accessed June 21, 2010].

Long, D. et al., 1999. *A Method for Forecasting Commercial Air Traffic Schedule in the Future*, McLean, Virginia: NASA. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.2256&rep=rep1&type=pdf [Accessed April 15, 2010].

Mackenzie-Williams, P., 2005. Aviation benchmarking: Issues and industry insights from benchmarking results. *Benchmarking: An International Journal*, 12(2), 112 - 124.

Mann, R., Adebanjo, O. & Kehoe, D., 1999. Best practices in the food and drinks industry. *British Food Journal*, 101, 238-254.

Marchand, M.G. et al., 1984. *The Performance of public enterprises*, North-Holland.

Martin, J. & Voltes-Dorta, A., 2008. International Airports: Economies of scale and marginal costs. *Journal of the Transportation Research Forum*, 47(1), 5-22.

Martín, J.C. & Román, C., 2001. An application of DEA to measure the efficiency of Spanish airports prior to privatization. *Journal of Air Transport Management*, 7(3), 149-157.

Matt Erskine, Greater Washington Board of Trade, 2009. Telephone call between Matt Erskine, Greater Washington Board of Trade, and David Schaar, George Mason University.

Matzko, M. & Wingfield, C., 1995. Coming: a source of competitive benchmarking for retail distribution strategy. *Journal of Retail Banking*, 17(2), 9.

McNair, C. & Leibfried, K., 1992. *Benchmarking, A Tool for Continuous Improvement*, New York, N.Y.: Harper Business Press.

Mehrabian, S. et al., 2000. An Assurance Interval for the Non-Archimedean Epsilon in DEA Models. *Operations Research*, 48(2), 344-347.

Metropolitan Washington Airports Authority, 2009. Metropolitan Washington Airports Authority, Facts. Available at: http://www.mwaa.com/about_the_authority/facts [Accessed December 9, 2009].

Michael Cintron, International Air Passengers' Association, 2009. Telephone call between Michael Cintron, International Air Passengers' Association and David Schaar, George Mason University.

Min, H. & Min, H., 1996. Competitive benchmarking of Korean luxury hotels using the analytic hierarchy process and competitive gap analysis. *Journal of Services Marketing*, 10, 58-72.

Mitchell, R.K., Agle, B.R. & Wood, D.J., 1997. Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *The Academy of Management Review*, 22(4), 853-886.

Morrison, W.G., 2009. Understanding the complexities and challenges of airport performance benchmarking. *Journal of Airport Management*, 3(2), 145 - 158.

Neufville, R.D. & Odoni, A.R., 2003. *Airport systems*, McGraw-Hill Professional.

Nyshadham, E.A. & Rao, V.K., 2000. Assessing Efficiency of European Airports: A Total Factor Productivity Approach. *Public Works Management Policy*, 5(2), 106-114.

Offerman, H., Simulation to support the airport stakeholder decision-making process. *Air & Space Europe*, 3(1-2), 60-67.

Office of Management and Budget, 2010. OIRA Statistical Programs and Standards.

OIRA Statistical Programs and Standards. Available at: http://www.whitehouse.gov/omb/inforeg_statpolicy/#ms [Accessed March 17, 2010].

Oum, T.H. & Yu, C., 2004. Measuring airports' operating efficiency: a summary of the 2003 ATRS global airport benchmarking report. *Transportation Research Part E: Logistics and Transportation Review*, 40(6), 515-532.

Pacheco, R.R. & Fernandes, E., 2003. Managerial efficiency of Brazilian airports. *Transportation Research Part A: Policy and Practice*, 37(8), 667-680.

Parker, D., 1999. The Performance of BAA before and after Privatisation: A DEA Study. *Journal of Transport Economics and Policy*, 33(2), 133-145.

Pat Oldfield, United Airlines, 2009. Telephone call between Pat Oldfield, United Airlines and David Schaar, George Mason University.

Paul McKnight, Jacobs Consultancy, 2009. Telephone call between Paul McKnight, Jacobs Consultancy, and David Schaar, George Mason University.

Pels, E., Nijkamp, P. & Rietveld, P., 2001. Relative efficiency of European airports. *Transport Policy*, 8(3), 183-192.

Performance Improvement Management, 2010. Performance Improvement Management DEA Software. Available at: http://www.deasoftware.co.uk/ [Accessed June 1, 2010].

Peter Stettler, Ricondo and Associates, 2009. Telephone call between Peter Stettler, Ricondo and Associates, and David Schaar, George Mason University.

Pezzimenti, J. & Macdonald, L., 2005. Calculating Airport Debt Service Coverage.

Plunkett, J.W., 2007. *Plunkett's Airline, Hotel & Travel Industry Almanac 2008*, Plunkett Research, Ltd.

Port of Seattle, Sea-Tac Airport, 2007. *Port of Seattle,Seattle-Tacoma International Airport,Greenhouse Gas Emissions Inventory - 2006*,

Public Utility Research Center, University of Florida, 2008. Regulation Body of

Knowledge. Available at: http://www.regulationbodyofknowledge.org [Accessed January 27, 2010].

Ray, S.C., 2004. *Data envelopment analysis*, Cambridge University Press.

Rees, J., 1986. *Technology, regions, and policy*, Rowman & Littlefield.

Rhoades, D.L., Jr, B.W. & Young, S., 2000. Developing a quality index for US airports. *Managing Service Quality*, 10(4), 257 - 262.

Roberts, F.S., 1972. What if utility functions do not exist? *Theory and Decision*, 3(2), 126-139.

Rollenhagen, M., 2003. Airport Expansion. *The Plain Dealer*. Available at: http://www.cleveland.com/indepth/airport/index.ssf?/indepth/airport/more/10532504655132.html [Accessed March 22, 2010].

Rousseau, J.J. & Semple, J.H., 1995. Radii of Classification Preservation in Data Envelopment Analysis: A Case Study of 'Program Follow-Through'. *The Journal of the Operational Research Society*, 46(8), 943-957.

San Diego International Airport, 2010. Facts & Statistics at San Diego International Airport. Available at: http://www.san.org/sdia/at_the_airport/education/airport_statistics.aspx [Accessed March 17, 2010].

Saneifard, R. et al., 2007. Euclidean Ranking DMUs with Fuzzy Data in DEA. *Applied Mathematical Sciences*, 1(60), 2989-2998.

Sarkis, J., 2000. An analysis of the operational efficiency of major airports in the United States. *Journal of Operations Management*, 18(3), 335-351.

Sarkis, J. & Talluri, S., 2004. Performance based clustering for benchmarking of US airports. *Transportation Research Part A: Policy and Practice*, 38(5), 329-346.

Schaar, D. & Sherry, L., 2010. Analysis of Airport Stakeholders. In *2010 Integrated Communications Navigation and Surveillance (ICNS) Conference*. 2010 Integrated Communications Navigation and Surveillance (ICNS) Conference. Herndon, VA.

Schaar, D. & Sherry, L., 2008. Comparison of Data Envelopment Analysis Methods Used in Airport Benchmarking. In 3rd International Conference on Research in Air Transportation – ICRAT 2008. Fairfax, VA. Available at: 3rd International Conference on Research in Air Transportation – ICRAT 2008 [Accessed August 4, 2009].

Schap, D., 1986. *Municipal ownership in the electric utility industry*, Praeger.

Schumer, C., 2008. *"Flight Delays Cost Passengers, Airlines and the U.S. Economy Billions". A Report by the Joint Committee Majority Staff*, Available at: http://jec.senate.gov/public/index.cfm?p=FlightDelaysCostPassengers [Accessed June 10, 2010].

SEIU USW West, 2009. SEIU USW West - About Us. Available at: http://www.seiu-usww.org/about/Default.aspx [Accessed December 16, 2009].

Seth Lehman and Emma Walker, Fitch Ratings, 2009. Telephone call between Seth Lehman and Emma Walker, Fitch Ratings, and David Schaar, George Mason University.

Smith, D.M., 1979. *Where the grass is greener*, Taylor & Francis.

Spendolini, M., 1992. *The Benchmarking Book*, New York, N.Y.: Amacom.

The Boeing Company, 2010. Boeing: Commercial Airplanes - Products. *Boeing: Commercial Airplanes - Products*. Available at: [Accessed March 22, 2010].

The Port Authority of New York and New Jersey, 2009. Overview of Facilities and Services - About the Port Authority. Available at: http://www.panynj.gov/about/facilities-services.html [Accessed December 9, 2009].

Thompson, I. & Cox, A., 1997. Don't imitate, innovate. *Dupply Management*, 2(22), 40-42.

Tobin, J., 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1), 24-36.

Tone, K., 2001. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3), 498-509.

Tracy, J.A., 2009. *How to Read a Financial Report: Wringing Vital Signs Out of the Numbers* 7th ed., Wiley.

Transportation Security Administration, 2009. Transportation Security Administration - What We Do. Available at: http://www.tsa.gov/what_we_do/index.shtm [Accessed December 10, 2009].

Tretheway, M.W., II, W.G.W. & Fok, A.K., 1997. The Total Factor Productivity of the Canadian Railways, 1956-91. *Journal of Transport Economics and Policy*, 31(1), 93-113.

Tulkens, H., 1993. On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. *Journal of Productivity Analysis*, 4(1), 183-210.

U.S. Census Bureau, 2008. *Methodology for the State and County Total Resident Population Estimates*, U.S. Census Bureau. Available at: http://www.census.gov/popest/topics/methodology/2008-st-co-meth.pdf.

U.S. Census Bureau, 2002. *NAICS 2002 definitions*, Available at: http://www.census.gov/epcd/naics02/naics02definitions.txt [Accessed July 17, 2010].

U.S. Census Bureau, 2010a. North American Industry Classification System. *North American Industry Classification System*. Available at: http://www.census.gov/eos/www/naics/ [Accessed March 26, 2010].

U.S. Census Bureau, 2010b. Population Estimates. Available at: http://www.census.gov/popest/metro/metro.html [Accessed March 17, 2010].

Ulusoy, G. et al., 2001. Benchmarking best manufacturing practices: A study into four sectors of Turkish industry. *International Journal of Operations & Production Management*, 21, 1020-1043.

Unite Here, 2009. Unite Here. Available at: http://www.unitehere.org/about/airports.php [Accessed December 16, 2009].

United States Code of Federal Regulations, 2010. Title 14, Part 91. Available at: [Accessed February 5, 2010].

Upham, P.J., 2003. *Towards sustainable aviation*, Earthscan.

US-Citizens Aviation Watch, 2009. US-Citizens Aviation Watch. Available at: http://www.us-caw.org/ [Accessed December 16, 2009].

Wang, D. & Sherry, L., 2007. Trend Analysis of Airline Passenger Trip Delays. In 2007 TRB 86th Annual Meeting. Washington, DC.

Wells, A.T. & Young, S.B., 2003. *Airport planning & management*, McGraw-Hill Professional.

White, E.T., 1976. *Utilities*, Dept. of Defense, National Defense University, Industrial College of the Armed Forces.

Wolfe, H.P. & NewMyer, D.A., 1985. *Aviation industry regulation*, SIU Press.

Yasin M. M., 2002. The theory and practice of benchmarking: then and now. *Benchmarking: An International Journal*, 9, 217-243.

Zairi, M., 1998. Benchmarking at TNT Express. *Benchmarking: An International Journal*, 5, 138-149.

Zhu, J. & Cook, W.D., 2007. *Modeling data irregularities and structural complexities in data envelopment analysis*, Springer.

Curriculum Vitae

David Schaar earned his M. Sc. degree in Industrial Engineering and Management from the Linköping Institute of Technology, Sweden, in 2003. His professional experience includes work as a Project Manager at the Corporate Executive Board where he performed benchmarking of the performance of I.T. departments at Fortune 500 corporations, and work at Metron Aviation developing methods for optimization of the use of enroute airspace.